

1998

A System Approach to the Design of Multirate Filter Banks.

Jian Huang

Louisiana State University and Agricultural & Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_disstheses

Recommended Citation

Huang, Jian, "A System Approach to the Design of Multirate Filter Banks." (1998). *LSU Historical Dissertations and Theses*. 6740.
https://digitalcommons.lsu.edu/gradschool_disstheses/6740

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Historical Dissertations and Theses by an authorized administrator of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

A SYSTEM APPROACH TO THE DESIGN OF MULTIRATE FILTER BANKS

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Electrical and Computer Engineering

by

Jian Huang

B.S.E.E., Shanghai University, China, 1990

M.S.E.E., Louisiana State University, 1994

August 1998

UMI Number: 9902642

UMI Microform 9902642
Copyright 1998, by UMI Company. All rights reserved.

**This microform edition is protected against unauthorized
copying under Title 17, United States Code.**

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

ACKNOWLEDGMENTS

I wish to express my appreciation to Dr. Guoxiang Gu, my major professor, for his academic guidance and support during my program. He was the first person who introduced me to the multirate digital signal processing, a technique that I have come to love and about which I wish to continue to learn even more. I also wish to thank him for his understanding of my responsibilities other than study and work, and for his patience and encouragement during the several long years of my study. He is not only my major professor, but also my life-long mentor.

Appreciation is also extended to my graduate committee, Dr. Jorge L. Aravena, Dr. Mark G. Davidson, Dr. Fred I. Denny, Dr. Sitharama S. Iyengar, and Dr. Kemin Zhou, for their guidance and review of this dissertation. The financial support from the Department of Electrical and Computer Engineering at LSU, Army Research Office, and Air Force Office for Scientific Research is also acknowledged.

I want to express my deep love to my parents, who always valued education, and who instilled that value in me. No matter under what circumstance, they have always had words to excite me to face, and sometimes, to overcome the challenges of life. Finally, I am grateful to the God for having given me so much.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	ii
LIST OF FIGURES	v
SPECIAL NOTATIONS	vii
ABSTRACT	viii
CHAPTER	
1 INTRODUCTION	1
1.1 Historical Overview	1
1.2 Dissertation Contributions	8
1.3 Background Materials	12
1.3.1 Basics of Multirate Signal Processing	12
1.3.2 Basics in Linear System Theory	16
2 CONVERGENCE RESULTS ON THE DESIGN OF QMF BANKS	20
2.1 Introduction	20
2.2 Reformulation of the Design Problem	21
2.3 An Iterative Algorithm and Its Convergence Analysis	30
2.4 A Different Iterative Algorithm and Its Convergence Analysis	36
3 TWO-CHANNEL QMF BANK DESIGN	39
3.1 Summary	39
3.2 Design of Synthesis Filter Banks with \mathcal{H}_2 and \mathcal{H}_∞ Optimizations	40
3.2.1 Formulation of the Design Approach	41
3.2.2 Optimal \mathcal{H}_2 and \mathcal{H}_∞ Solutions for the Synthesis Filter Bank	45
3.2.3 Design Algorithms for the Synthesis Filter Bank	49
3.3 State-space Computation	50
3.3.1 State-space Realizations	50
3.3.2 Inner-outer Factorization	52
3.3.3 Computation of \mathcal{H}_2 Norm	55
3.3.4 Computation of Hankel Norm	56
3.4 Design Examples	58
4 MULTI-CHANNEL FILTER BANK DESIGN AND ANALYSIS	65
4.1 Multirate Systems to be Designed	65
4.2 Design of Synthesis Filter Banks with \mathcal{H}_2 and \mathcal{H}_∞ Optimizations	66
4.2.1 Formulation of the Design Approach	66

4.2.2	Optimal \mathcal{H}_2 and \mathcal{H}_∞ Solutions for the Synthesis Filter Bank .	74
4.2.3	Design Algorithms for the Synthesis Filter Bank	75
4.3	State-space Computation	76
4.4	An Illustrative Example	79
4.5	Extension to Nonuniform Filter Banks	83
4.5.1	A Two-Channel Nonuniform Filter Bank	85
4.5.2	Multi-Channel Nonuniform Filter Banks	86
5	DESIGN OF 2-D QMF BANKS	98
5.1	Introduction	98
5.2	Error Index Function and 2-D Linear Phase Filters	99
5.3	Computation of the Stop Band Error E_s	102
5.4	Computation of the Reconstruction Error E_r	109
5.5	An Iterative Algorithm and Its Convergence Analysis	115
5.6	Computation of Integral Error	117
6	CONCLUSIONS	124
	BIBLIOGRAPHY	127
	VITA	136

LIST OF FIGURES

1.1	The M -fold decimator and L -fold interpolator.	13
1.2	Two interconnections of decimators with expanders.	13
1.3	The noble identities for multirate systems.	13
1.4	Sampling rate alteration systems	13
1.5	The polyphase identity, a special cascade and its equivalent scheme. . . .	13
2.1	Two-channel QMF bank.	22
3.1	Two-channel error system.	43
3.2	Example 3.1: $ H_0 $ and $ H_1 $ in dB vs. $\omega/2\pi$	62
3.3	Example 3.1: $ \hat{F}_0 $ and $ \hat{F}_1 $ in dB vs. $\omega/2\pi$	62
3.4	Example 3.1: Reconstruction and aliasing error vs. $\omega/2\pi$	63
3.5	Example 3.2: $ H_0 $ and $ H_1 $ in dB vs. $\omega/2\pi$	63
3.6	Example 3.2: $ \hat{F}_0 $ and $ \hat{F}_1 $ in dB vs. $\omega/2\pi$	64
3.7	Example 3.2: Reconstruction and aliasing error vs. $\omega/2\pi$	64
4.1	M -channel maximally decimated filter bank.	67
4.2	Multi-channel error system.	67
4.3	Example 4.1: $ H_0 $, $ H_1 $, and $ H_2 $ in dB vs. ω/π	81
4.4	Example 4.1: $ \hat{F}_0 $, $ \hat{F}_1 $, and $ \hat{F}_2 $ in dB vs. ω/π	81
4.5	Example 4.1: Reconstruction error vs. $1000\omega/\pi$	82
4.6	Example 4.1: Aliasing error	82
4.7	The general multirate filter system.	84

4.8	A two-channel nonuniform system.	87
4.9	Equivalent system of an analysis filter bank branch.	87
4.10	Equivalent system of a synthesis filter bank branch.	88
4.11	An equivalent system of a 2×2 identity matrix.	88
4.12	Two equivalent systems.	89
4.13	An equivalent uniform system to Figure 4.8.	89
4.14	Kovacevic's Transform 1.	94
4.15	Desired synthesis bank relating to Figure 4-14, $S'_i(z) = z^{-n_i} S_i(z)$	94
4.16	Desired uniform system, $S'_i(z) = z^{-n_i} S_i(z)$	95
4.17	Equivalent system to Figure 4.15.	95
4.18	Same analysis bank as in Figure 4.14, $A'_i(z) = z^{m_i} A_i(z)$	96
5.1	A 2-D two-channel multirate filter bank.	99

SPECIAL NOTATIONS

A^* :	Complex conjugate transpose of A
A^T :	Transpose of A
A^{-1} :	Inverse of A
AC:	Alias component
ARE:	Algebraic Riccati equation
\mathbf{C} :	Field of complex numbers
FIR:	Finite impulse response
\mathcal{H}_2 :	Hilbert space of square integrable functions on a unit circle which admit an analytic extension to outside of unit circle.
\mathcal{H}_∞ :	Banach space of essentially bounded functions on a unit circle which admit an analytic extension to outside of unit circle.
IIR:	Infinite impulse response
PR:	Perfect reconstruction
QMF:	Quadrature-mirror-filter
$\tilde{Q}(z)$:	$= Q^T(z^{-1})$ (discrete-time)
\mathbf{R} :	Field of real number
λ :	Lagrange multiplier, constant
$\sigma_{\max}(A)$:	Largest singular value of A
$\ \bullet\ _\infty$:	\mathcal{H}_∞ norm
$\ \bullet\ _H$:	Hankel norm

ABSTRACT

This dissertation studies the design of multirate filter banks by adopting a so-called system approach. The design issue of Johnston's method is first investigated in which an explicit expression of the reconstruction error is derived using Lyapunov stability theory, and new convergent iterative algorithms are proposed through non-linear optimization. The results are extended to the two-dimensional filter banks. The design issue of more general multirate filter banks is also investigated through model matching method. Using standard results from modern control theory, new design algorithms are developed which minimize the reconstruction error while completely eliminating the aliasing error. State-space realizations, inner-outer factorizations, and optimal Hankel norm approximation are used to reduce the complexity of computation and improve the accuracy of the proposed design algorithms.

CHAPTER 1

INTRODUCTION

1.1 HISTORICAL OVERVIEW

The conventional digital signal processing structures belong to the class of single-rate systems since the sampling rates at the input and at the output and at all internal nodes are the same. Multirate digital signal processing systems employ two sampling rate alteration devices, the down-samplers and the up-samplers. Discrete-time systems with unequal sampling rates at various parts of the system are called multirate systems. The fact that a signal can be split into more than one channels using non-ideal filters, down-sampled and then reconstructed without spectral overlap (aliasing) was observed two decades ago [9]. Since then, the field of multirate filter banks has been growing considerably. The discipline of multirate digital signal processing techniques finds applications in speech and image compression, the digital audio industry, statistical and adaptive signal processing, numerical solution of differential equations, and in many other fields [90]. It also fits naturally with certain special classes of time-frequency representations such as the short-time Fourier transform and the wavelet transform, which are useful in analyzing the time-varying nature of signal spectra [52, 69, 90, 95].

Typically, for example, in digital audio field, three different sampling rates are presently adopted: 32 kHz in broadcasting, 44.1 kHz in digital compact disk (CD), and 48 kHz in digital audio tape (DAT) [44]. Conversion of sampling rates of audio signals among these three different rates is often necessary in many situations. Because the sampling space in a multirate system can vary from point to point,

the sampling rates at various internal points can be kept as small as possible. This leads to more efficient processing of signals. Unfortunately multirate systems also result in a new type of error, aliasing, which should be cancelled completely. Early work in multirate signal processing has focused on aliasing cancellation. An interesting development is that after the initial results on aliasing cancellation, Mintzer, Smith and Barnwell independently demonstrated the possibility of obtaining the exact reconstruction of the input signal at the output [53, 77]. Applications to speech and image compressions in subband coding are the impetus to research of multirate filter banks. In such filter banks, a digital signal is filtered and split into different frequency bands which is then down-sampled, coded, and transmitted. At the receiver end, the transmitted signals can then be decoded, filtered, up-sampled, and the original signal recovered by the perfect reconstruction. Thus the subband coding technique offers great efficiency in compression of the digital data, and multirate filter banks are ideal for providing the tool to accomplish data compressions. The last fifteen years or so have witnessed the rapid development in the design of multirate filter banks. Many new design techniques are proposed, and improved which are widely used in engineering applications. Moreover perfect reconstruction has been extensively studied from two-channel case to multi-channel ones by many researchers [17, 28, 40, 58, 78, 86, 90, 87, 89, 93], including the detailed analysis of lossless (orthogonal) systems and the building of linear phase filter banks. Comprehensive overviews of this subject can be found in [90, 94].

Multirate filter banks can be divided into uniform filter banks and nonuniform ones. The sampling rates for uniform filter banks are not only integers but also the same. This leads to a uniform division of the signal spectrum. Most research falls into this category. Without specifically mentioned, “filter bank” usually means

uniform filter bank. On the other hand, a nonuniform filter bank has unequal sampling rates which can be rational, i.e., non-integer. This leads to a nonuniform division of the signal spectrum. Hence they are more general and versatile because of the flexibility in their structure.

While nonuniform multi-channel filter banks are most interesting and challenging, the initial study in multirate signal processing focused on two-channel multirate filter banks consisting of half band filters. These are the simplest filter banks. But if the underlying filters are linear phase, then perfect reconstruction is not trivial at all ¹. In fact it is known that a sufficient condition of power complementary [90] for perfect reconstruction holds for only those linear phase filters with trivial magnitude response. For this reason Johnston proposed a simple optimization method for the design of two-channel linear phase multirate filter banks through minimization of reconstruction error and stop band error in magnitude response [36]. This is the earliest design method which results in not only high performance filter banks, but also efficient design procedures that are easy to understand and to implement. However it is difficult to be extended to multi-channel multirate filter banks, and it does not admit property of perfect reconstruction. A new paradigm is needed in order to tackle the general design issues and meanwhile minimize or even eliminate the reconstruction error. This gives the rise of Vaidyanathan's work on perfect reconstruction of general uniform multirate filter banks by using paraunitary and lossless system theory which provides a unified framework for multi-channel filter banks [90]. By adopting (lossless) polyphase matrices, Vaidyanathan showed how perfect reconstruction can be accomplished by restricting a polyphase component

¹Linear phase filters have the property of preserving the waveform of the input signals which are preferred in engineering applications.

matrix in the analysis bank to be FIR and lossless based on lattice structure decomposition of the filter bank. This is equivalent to enforce the alias component matrix to be FIR and lossless. A lossless alias component matrix implies that the analysis filters satisfy the power-complementary property. Therefore, the aim is to optimize the parameters in the decomposed lattice filters so that the sum of the stopband energies of the analysis filters is minimized. The advantage of the new nonlinear optimization is perfect reconstruction. The price paid is the increase of computation complexity with the drawback of nonlinear phase. The problem of nonlinear phase is later resolved in [62, 63] by giving up the power complementary condition. But the resulting filters have relatively high orders compared with Johnston's method.

A notable work parallel to Vaidyanathan's is that of Vetterli [93]. He independently introduced the concepts of alias component matrix and polyphase component matrix in a different way and in different terminology such as modulated filter matrix and polyphase filter matrix. By using matrix properties and matrix operation, the conditions for aliasing error and amplitude distortion elimination were given. Also, a general sufficient condition and a specific necessary condition for perfect reconstruction were derived. Moreover, Vetterli linked the relation between filter banks and transmultiplexers. Furthermore, when concerning the synthesis filters, he showed that the design was a trade-off between individual filter quality, reconstruction quality, and input-output delay.

In contrast to Johnston, Vaidyanathan, and Vetterli, the work of Nayebi, Smith, and Barnwell took a different approach in which the analysis/synthesis system was analyzed in the time domain [58]. Every filter bank system was divided into three sections, the analysis section, the down-up sampling section, and the synthesis section. For each of those three sections, an input/output relationship was derived

in vector form. All analysis and synthesis filters were treated as having the same length which is an integer multiple of the decimation rate. Through matrix manipulation, the input/output relationship of the filter bank was then expressed in a general block-matrix equation. This led to a set of necessary and sufficient conditions for exact (perfect) reconstruction. Although the general block-matrix equation is overconstrained so that it is not solvable for nontrivial cases, given the analysis filters, it is always possible to find the optimum synthesis filters by minimizing a feasible norm. The design algorithms are based on an iterative process in which the reconstruction error is reduced to a desired value which can be set to zero, equivalent to perfect reconstruction. Through this time-domain approach, Nayebi, Smith, and Barnwell obtained several satisfactory results in multirate filter bank design [59, 60, 61].

The success in perfect reconstruction for uniform filter banks motivated research for nonuniform filter banks. Hoang and Vaidyanathan considered the simplest type of nonuniform filter banks consisting of integer sampling rates [26]. Necessary conditions for elimination of aliasing error were derived, and a specific example was used to demonstrate the possibility of perfect reconstruction. It was achieved through conversion of nonuniform filter banks into uniform ones. Their method was later classified as “indirect”. A direct method was proposed afterwards by Kovacevic and Vetterli [43] in conversion of nonuniform filter banks into uniform ones which has less computation complexity leading to better design in terms of filter responses. More importantly the direct method applies to those nonuniform filter banks with rational sampling rates. However it should be mentioned that the direct method can not replace the indirect method completely in the sense that there exist cases where the direct method does not work, while the indirect method works. Perhaps

a more serious issue for nonuniform filter banks is the so called “structure dependency” which until today is still an open problem. Other work in nonuniform filter banks can be found in the literature which includes time domain approach [59], and lifting or blocking technique [4].

Although perfect reconstruction is desirable, and is theoretically interesting, it may not be practical in engineering applications. Recall that the design is a trade-off between various qualities of individual filters, reconstruction error, and input-output delay. Therefore, in using linear phase filters, there is a direct conflict between qualities of magnitude response of individual filters and signal reconstruction. Design methods based on perfect reconstruction usually have higher computation complexity, and higher cost in implementation. Considering that other design methods often generate near-perfect reconstruction, it is really not necessarily to enforce zero reconstruction error at all cost. In fact exact reconstruction does not exist in reality because of the existence of the coding error. Linear phase and free of aliasing are much more important features for preserving the original signal. Other considerations are relatively secondary. For this reason, a new line of research emerged recently [74, 3, 4]. It separates the design of analysis filter bank from that of synthesis bank which allows more freedom. Indeed, analysis banks can now be designed independently solely for subband coding. Synthesis banks are then designed through minimization of the reconstruction error which serves better the purpose of recovering the original signal. The advantage of this new design method lies in the fact that analysis and synthesis filter banks have different roles in multirate signal processing. Different treatments to different filter banks improve the overall performance of the filter bank, including coding quality, and simplify the design procedures as well. It was Shenoy, Burnside, and Parks who proposed first the model matching concept

that inspired other work in the same problem area. The most interesting ones are [3, 4] in which Chen and Francis investigated multirate filter banks by using blocking (also known as lifting) technique as well as \mathcal{H}_∞ optimization. One unique feature of this design method is the usage of theory from linear control systems. The model matching error is measured by l_2 -induced norm which is the same as \mathcal{H}_∞ norm of the error system. Given causal and stable FIR or IIR analysis filters, and given a tolerable time delay, their work gives a systematic design procedure to design causal and stable IIR synthesis filters to minimize the \mathcal{H}_∞ norm, for which there is a ready-made software. The limitation of this design method is that in addition to nonzero aliasing error, quite high order synthesis filters may be required in order to achieve good reconstruction error due to unblocking operation. Therefore, model reduction was employed to reduce the order of synthesis filters in [3].

Before concluding this section, it should be mentioned that the success of sub-band coding also encouraged researchers to extend the ideas to multidimensional signals. Because of applications in image compression and coding, two dimensional multirate systems have attracted great attention. Vetterli [92] first demonstrated the possibility of using the multirate filter banks in multiple dimensions. Since then progress has been made by many researchers of which a natural approach explored most is to obtain multidimensional filters from their one-dimensional counterparts. Among all the different methods, the one demonstrated by Chen and Vaidyanathan [6, 7] is probably the most general and effective [38]. According to their study, an one dimensional low pass prototype filter can be designed first, a separable multidimensional filter can then be constructed from the prototype filter, and finally the desired filter bank can be obtained by matrix decimation. This method works for arbitrary dimensions and arbitrary parallelepiped-shaped passbands. The Nyquist

constraint and zero-phase (known as linear phase in 1-D case) requirement can both be satisfied. Moreover, the polyphase components all have separable denominators, although the initial multidimensional filters may be nonseparable. The drawback of the design method in [6, 7] is the dependence on 1-D method. Because the set of all multi-dimensional filters is much greater than the set of transformed ones, the optimality of the design method in [6, 7] is questionable. Despite of many difficulties, research results in multidimensional filter banks grow at a rapid rate. However other existing design methods are less successful, and thus will not be covered in this section.

1.2 DISSERTATION CONTRIBUTIONS

This dissertation is a continuation of the existing research in multirate digital signal processing, emphasizing on the design of multirate filter banks for the purpose of data compression and subband coding. In spite of tremendous progress in this growing area, many problems remained of which some will be treated in this work. Specifically the model matching method proposed in [3] will be pursued, and re-examined. As mentioned early, a drawback of the results in [3] is the existence of the aliasing error. Although small reconstruction error implies small aliasing error, the actual reconstruction error is very sensitive to the aliasing error because of the existence of the coding error. How to minimize the reconstruction error while eliminating the aliasing error is thus an open problem which will be studied in this dissertation with a satisfactory answer. Another problem is the design issue of Johnston for two-channel filter banks [36]. Thus far there exist a number of practical iterative algorithms which lack convergence analysis. This problem will be investigated for both one-, and two-dimensional multirate filter banks for which

new algorithms are proposed, and convergence issue resolved. An important feature of this dissertation is the use of modern control theory for linear time-invariant systems. System theories such as Lyapunov stability and optimal control are the main tools. Indeed by using Lyapunov method, an explicit expression of the reconstruction error is derived for the two channel filter banks studied by Johnston [36] which allows further analysis of the convergence of the proposed iterative algorithms. Using optimal control theory, a design algorithm is developed which minimizes the reconstruction error while eliminating the aliasing error completely. State-space realization and inner-outer factorization are used as well to reduce the computation load to improve the efficiency of the proposed design algorithm.

The contributions of this dissertation are stated in more detail as follows:

- Development of new iterative algorithms for the design problem of Johnston [36] which are proven to be convergent.

The design problem of Johnston for two-channel filter banks [36] will be investigated in Chapter 2 which considers the design of quadrature mirror filter (QMF) banks whose analysis and synthesis filters are FIR and have linear phase. The design criterion is of least-square type with an error index which is a linear combination of the reconstruction error and the stop band error in quadratic norm [36]. Such an error index is highly nonlinear and its global minimum is difficult to compute in finite algorithmic steps. Thus several iterative algorithms have been previously developed for minimizing this particular nonlinear error index function [2, 21, 34, 70]. However there lacks convergence analysis on these iterative algorithms in the literature. This dissertation proposes a new algorithm that is a modification of the one developed in [2] and proves its convergence. The results include a new derivation for an explicit

expression of the error index to be minimized and a necessary condition for minimality. These results offer new insight to the design of QMF banks and relate it to a more general nonlinear optimization problem. Moreover it is shown that this new algorithm is a quasi-Newton descend algorithm, and thus it not only converges, but also admits fast convergence rate in the vicinity of the minimum solution. The same method is applied to the iterative algorithm developed in [34] as well which yields similar convergence results.

- Development of an efficient design algorithm that minimizes model matching error for QMF banks subject to zero aliasing error.

The design of quadrature mirror filter banks is studied again in Chapter 3 via frequency domain optimizations. A new approach is adopted that gives the necessary and sufficient condition for perfect reconstruction (PR). While analysis filter banks are designed to achieve frequency domain specifications required for subband coding, synthesis filter banks are designed to minimize the reconstruction error in frequency domain that achieves zero aliasing error. Two criteria are used to measure the reconstruction error: \mathcal{H}_2 or RMS (root-mean-square), and \mathcal{H}_∞ or Chebyshev (sup-norm). All stable and causal synthesis filter banks that satisfy zero aliasing error are first parameterized, and state-space solutions are then derived for both \mathcal{H}_2 and \mathcal{H}_∞ optimizations. Efficient numerical algorithms are proposed to obtain the optimal synthesis filter bank. Moreover the asymptotic PR property is established for optimal \mathcal{H}_2 and \mathcal{H}_∞ solutions of the synthesis filter bank.

- Extension of the efficient design algorithm to multi-channel filter banks that minimizes reconstruction error subject to zero aliasing error.

For multi-channel filter banks, parameterization of all causal stable synthesis filter banks with zero aliasing error is more complicated. The reason is that the determinant of the aliasing component matrix can be complex valued while coefficients of the filter banks are real. This problem is resolved in Chapter 4 through the use of polyphase decomposition. The efficient and reliable algorithm is generalized from two-channel to multi-channel filter banks. Despite that the computation complexity and filter order increase as the number of channels grows, the care is taken to lower both so that the algorithm is comparable to the two-channel case. Moreover the conversion of nonuniform filter banks to uniform ones is also considered. Although [43] introduced a transform of the analysis filter bank into an equivalent uniform system, the transformation for the synthesis filter bank was not mentioned due to its complexity. This missing formulae is derived in this dissertation so that individual synthesis filters in the equivalent uniform system are exactly the polyphase components of the synthesis filters in the nonuniform filter banks. Therefore, this type of nonuniform multirate filter banks can now be easily designed by using the existing design techniques for the uniform multirate filter banks.

- Extension of the proposed iterative algorithm to two-dimensional two-channel multirate filter banks.

The design issue of two dimensional (2-D) two-channel multirate filter banks is raised in Chapter 5. The filter banks involve linear phase FIR diamond-shaped analysis and synthesis filters. The design criterion is of least-square type with an error index which is a linear combination of the reconstruction error and the stop band error in quadratic norm. Due to the limitation of the filter orders in two different direction, the reconstruction error in one

dimensional case can not be generalized directly. Modification is made so that the error index function is meaningful for such 2-D FIR filter banks. The results include a derivation of an explicit expression for the error index to be minimized, the necessary condition for minimality, and iterative algorithms to obtain minimal solutions. This is an extension of the results in Chapter 2 that was initially reported in [25] focusing on one dimensional case.

This dissertation is comprehensive in using the system theory to approach multirate digital signal processing. The results are submitted to journals of IEEE of which one appeared [25], one accepted [30], and others under review [31, 32]. It is hoped that the results in this dissertation will improve the technology for multirate digital signal processing.

1.3 BACKGROUND MATERIALS

This section introduces some most fundamental concepts of multirate signal processing and linear control systems. The integration of signal processing and system control enables us to adopt a system approach to the design of multirate filter banks.

1.3.1 BASICS OF MULTIRATE SIGNAL PROCESSING

Only the very basic concepts shown from Figure 1.1 to 1.5 will be introduced.

Decimation and Interpolation

Multirate signal processing deals with discrete-time sequences taken at different sampling rate. To achieve sampling rate conversion, two basic devices are used, one is called decimator (down-sampler, or compressor) and the other is called interpolator (up-sampler, or expander).

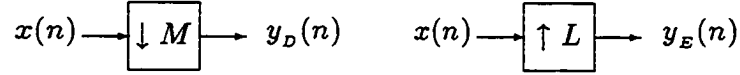
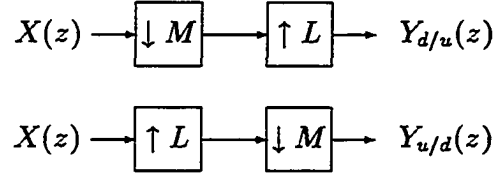
Figure 1.1: The M -fold decimator and L -fold interpolator.

Figure 1.2: Two interconnections of decimators with expanders.

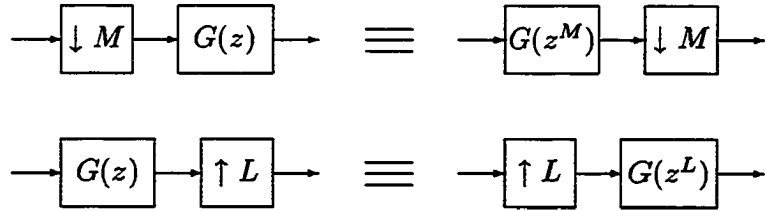


Figure 1.3: The noble identities for multirate systems.

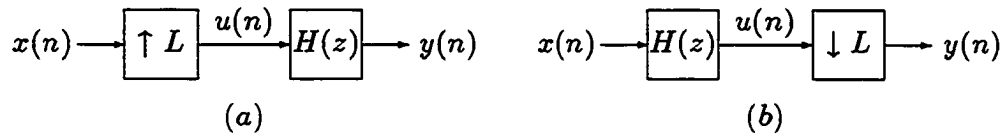


Figure 1.4: Sampling rate alteration systems: (a) interpolator and (b) decimator.

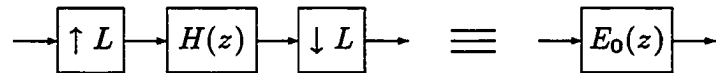


Figure 1.5: The polyphase identity, a special cascade and its equivalent scheme.

Figure 1.1 shows an M -fold decimator and an L -fold interpolator where M and L are integers. The M -fold decimator takes an input sequence $x(n)$ and produces an output sequence

$$y_D(n) = x(Mn).$$

In the \mathcal{Z} -transform domain,

$$Y_D(z) = \frac{1}{M} \sum_{k=0}^{M-1} X(z^{1/M} W_M^k), \quad W_M = e^{-j2\pi/M}. \quad (1.1)$$

Only those samples of $x(n)$ occurring at time equal to multiples of M are retained by the decimator. This results in a sequence $y_D(n)$ whose sampling rate is $1/M$ of that of $x(n)$. In general, it may not be possible to recover $x(n)$ from $y_D(n)$ because of loss of information due to aliasing. The L -fold interpolator takes an input $x(n)$ and produces an output sequence

$$y_E(n) = \begin{cases} x(n/L), & n = kL, k \in \mathbb{Z} \\ 0, & n \neq kL. \end{cases}$$

Note that the sampling rate of $y_E(n)$ is L times larger than that of the original sequence $x(n)$. In \mathcal{Z} -transform domain,

$$Y_E(z) = X(z^L). \quad (1.2)$$

The interpolator does not cause loss of information so that the input $x(n)$ can be recovered from $y_E(n)$ by L -fold decimation.

The interpolator and the decimator are linear but time-varying discrete-time systems. In practice, the zero-valued samples inserted by the interpolator are re-

placed with appropriate nonzero values using some type of interpolation process in order that the new higher rate sequence be useful. Both the interpolator and the decimator are often used together in a number of applications involving multirate signal processing. For example, an application of using both types of sampling rate alteration devices is to achieve a sampling rate change by a rational number rather than an integer value.

Cascade Equivalences

A complex multirate system is formed by an interconnection of the basic sampling rate alteration devices and the components of an linear time-invariant digital filter. In many applications, these devices appear in a cascade form. The basic sampling rate alteration devices can be used to change the sampling rate of a signal by an integer factor only. To implement a fractional change in the sampling rate, a cascade of a down-sampler and an up-sampler should be used. In Figure 1.2, using (1.1) and (1.2) for down and upsampling in z -domain,

$$Y_{u/d}(z) = \frac{1}{M} \sum_{k=0}^{M-1} X(z^{L/M} W_M^{kL}), \quad (1.3)$$

$$Y_{d/u}(z) = \frac{1}{M} \sum_{k=0}^{M-1} X(z^{L/M} W_M^k). \quad (1.4)$$

It is of interest to determine the condition under which such two cascades are equivalent. It can be shown that this equivalence is possible if and only if L and M are relatively prime (or coprime), i.e., L and M do not have a common factor that is an integer $k > 1$ [90].

Two other simple cascade equivalence relations are depicted in Figure 1.3. They are called noble identity 1 and 2, respectively. These rules enable us to move the

basic sampling rate alteration devices in multirate networks to more advantageous positions.

Sampling Rate Alteration Systems

Since up-sampling causes periodic repetition of the basic spectrum, the unwanted images in the spectra of the up-sampled signal must be removed by using a lowpass filter $H(z)$, called the interpolation filter, as indicated in Figure 1.4(a). On the other hand, as indicated in Figure 1.4(b), prior to down-sampling, the signal should be band limited to $|\omega| < \pi/M$ by means of a lowpass filter $H(z)$, called the decimation filter, to avoid aliasing caused by down-sampling. The system of Figure 1.4(a) is often simply called an interpolator while the system of Figure 1.4(b) is simply called a decimator.

A fractional change in the sampling rate can be achieved by cascading a factor-of- M decimator with a factor-of- L interpolator, where M and L are positive integers. Such a cascade is equivalent to a decimator with a decimation factor of M/L or, alternatively, to an interpolator with an interpolation factor of L/M . Such topic can be found in [66]. Figure 1.5 gives a special cascade with its equivalent scheme in which $E_0(z)$ is the 0th polyphase component of $H(z)$. [90] showed that if $H(z) = z^m$, then $E_0(z) = 1$ when $m = 0$. Otherwise, $E_0(z) = 0$.

1.3.2 BASICS IN LINEAR SYSTEM THEORY

A finite dimensional discrete linear time invariant system can be described by the following state-space equation:

$$x(k+1) = Ax(k) + Bu(k), \quad x(0) = x_0, \quad (1.5)$$

$$y(k) = Cx(k) + Du(k), \quad k \geq 0 \quad (1.6)$$

where $x(k) \in \mathbf{R}^n$ is called state variable, $x(0)$ initial condition, $u(k) \in \mathbf{R}^m$ is called the control input, and $y(k) \in \mathbf{R}^p$ the system output. The matrices A , B , C , and D are appropriately dimensioned real constant matrices. The corresponding transfer matrix from u to y is defined as

$$Y(z) = G(z)U(z)$$

where $U(z)$ and $Y(z)$ are the \mathcal{Z} -transform of $u(k)$ and $y(k)$ respectively with zero initial condition. Hence,

$$G(z) = C(zI - A)^{-1}B + D.$$

The four-tuple matrix (A, B, C, D) is a realization of $G(z)$. Several notions are important in linear system theory.

Definition 1.1 *The system described by state-space equation (1.5) or the pair (A, B) is said to be reachable if, for any initial state $x(0) = x_0$, $n > 0$ and final state x_1 , there exists an input $u(\cdot)$ such that the solution of (1.5) satisfies $x(n) = x_1$. Otherwise, the system or the pair (A, B) is said to be unreachable.*

Reachability is a system property, and independent of input and output of the system. A dual notion is the observability.

Definition 1.2 *The dynamical system described by the equation (1.5) and (1.6) or by the pair (C, A) is said to be observable if, for any $n > 0$, the initial state $x(0) = x_0$ can be determined from the time history of the input $u(k)$ and the output $y(k)$ in the interval of $[0, n]$. Otherwise, the system or the pair (C, A) is said to be unobservable.*

The duality between the above two notions is due to the equivalence of reachability of (A_1, B_1) and observability (C_2, A_2) if $A_1^T = A_2$, and $B_1^T = C_2$. The minimality of the state-space realization is hinged to the notions of reachability and observability.

Theorem 1.1 *A state-space realization (A, B, C, D) is minimal if, and only if (A, B) is reachable, and (C, A) is observable.*

If a state-space realization is not minimal, then Kalman decomposition [37] can be applied to eliminate the unreachable, or/and unobservable subsystem to obtain lower order state-space realization.

On the other hand, stability is also an important notion for system theory which is given in the next definition.

Definition 1.3 *An unforced dynamical system $x(k+1) = Ax(k)$ is said to be stable if for every nonzero initial condition $x(0) \in \mathbf{R}^n$, the state variable $x(k) \rightarrow 0$ asymptotically. This is equivalent to the condition that all eigenvalues of A lie strictly inside of unit circle. A system matrix A is called stable, if the corresponding unforced dynamic system is stable.*

Stabilizability and detectability are two parallel notions to reachability and observability respectively.

Definition 1.4 *The dynamical system (1.5), or the pair (A, B) , is said to be stabilizable if there exists a state feedback $u = Fx$ such that the overall system is stable, i.e., $A + BF$ is stable. The system, or the pair (C, A) , is said to be detectable if there exists an output injection $v = Ly$ such that the overall system is stable, i.e., $A + LC$ is stable.*

Stability of the dynamic system (1.5) can be determined by using the Lyapunov method.

Theorem 1.2 *(Lyapunov) Suppose that (C, A) is detectable. Then a matrix A is a stability matrix if and only if the Lyapunov equation*

$$A^T P A - P + C^T C = 0$$

admits a unique solution P which is positive definite. Similarly, suppose that (A, B) is stabilizable. Then a matrix A is a stability matrix if and only if the Lyapunov equation

$$A Q A^T - Q + B B^T = 0$$

admits a unique solution Q which is positive definite.

CHAPTER 2

CONVERGENCE RESULTS ON THE DESIGN OF QMF BANKS

2.1 INTRODUCTION

The design of QMF banks finds applications in subband coding and telecommunication systems [10, 14, 92]. Many design algorithms can be found in [90] and its references therein. One particular design problem is formulated as minimization of the linear combination of the reconstruction error for the QMF banks and stop band error for the prototype low pass filter, measured in quadratic norm [36]. Because the quadratic form of the reconstruction error is a fourth order function of the prototype filter coefficients, the minimization problem is a nonlinear one that is difficult to solve numerically. Several algorithms have been developed in the literature. The most notable ones are the iterative algorithms developed in [2, 34]. In [2], when designing the synthesis filters, the design method splits the 4th-order term in the objective function into a product of two squared terms. One of the two squared terms is fixed in order to obtain an objective function which is quadratic. In [34], at i th step, the analysis (or synthesis) filters are fixed, and the synthesis (or analysis) filters are designed to minimize the error index; then the synthesis (or analysis) filters are fixed, and analysis (or synthesis) filters are designed to minimize the error index. Because these two iterative algorithms have appealing property that the minimum solution at each iterative step is easy to compute. They become popular and have their importance in the design of QMF banks. Although these two algorithms are convergent for all the simulation examples, no convergence analysis

is given in the literature, and thus it is not clear whether or not the two iterative algorithms in [2, 34] are convergent in general.

This chapter considers the convergence issue for the iterative algorithms developed in [2, 34] on the design of QMF banks. The use of Lyapunov method enables us to express the reconstruction error in a compact matrix form. A new iterative algorithm is then proposed to search for the optimal solution together with the convergence analysis. In particular, minimization of the error index for the QMF design in [36] has been extended to a more general nonlinear optimization problem and a new iterative algorithm is proposed to solve this more general nonlinear optimization problem. The new proposed iterative algorithm, though motivated by the results in [2], is different from [2]. In fact the iterative algorithm developed in [2] is not applicable to the general nonlinear optimization problem considered in this dissertation. More importantly, it will be shown that the new proposed algorithm is a descend algorithm and is convergent. Its convergence is independent of the initial guess. Furthermore, the new proposed algorithm is a quasi-Newton algorithm and thus converges superlinearly in the vicinity of the local minimum [54]. Finally the iterative algorithm developed in [34] is also investigated and appropriate modifications are made. Its convergence is proven along the same line as the proof for the convergence of the proposed algorithm. It is emphasized that the convergence results in this chapter are most general and apply to other similar nonlinear optimization problems as well.

2.2 REFORMULATION OF THE DESIGN PROBLEM

The basic structure of a two-band QMF bank is shown in Figure 2.1. The signal $x(n)$ is split into two frequency subbands by two analysis filters $H_0(z)$ and $H_1(z)$

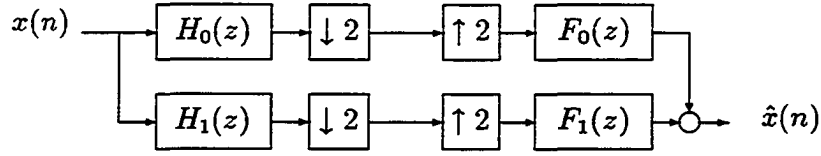


Figure 2.1: Two-channel QMF bank.

according to the energy distribution of the signal $x(n)$ in frequency domain. Each subband signal $x_i(n)$, $i = 1, 2$, is down-sampled by a factor of two, then coded and transmitted. We assume that the error caused by coding and transmission is negligible. In the receiver end, each subband signal $u_i(n)$ is up-sampled by a factor of two. After passing through two synthesis filters $F_0(z)$ and $F_1(z)$, respectively, the two subband signals, $\hat{x}_1(n)$ and $\hat{x}_2(n)$, are combined to form the reconstructed signal $\hat{x}(n)$. Because coding error has the least effect for linear phase filters, it is desirable for each of the filters in the QMF bank to have linear phase. The design objective is thus to find the two analysis filters $H_0(z)$, $H_1(z)$, and the two synthesis filters $F_0(z)$, $F_1(z)$, all FIR and linear phase, such that the reconstructed signal $\hat{x}(n)$ is approximate to the original signal $x(n)$ in some optimal sense. By Figure 2.1, the input/output relation in \mathcal{Z} -transform is given by

$$\begin{aligned} \hat{X}(z) &= \frac{1}{2} [H_0(z)F_0(z) + H_1(z)F_1(z)] X(z) \\ &\quad + \frac{1}{2} [H_0(-z)F_0(z) + H_1(-z)F_1(z)] X(-z). \end{aligned} \quad (2.1)$$

Conventionally, the synthesis filters $F_0(z)$ and $F_1(z)$ are chosen to be equal to the analysis filters $2H_1(-z)$ and $-2H_0(-z)$, respectively, to eliminate the aliasing error. Moreover by choosing $H_1(z) = H_0(-z)$, (2.1) is reduced to the form

$\hat{X}(z) = [H_0(z)H_0(z) - H_0(-z)H_0(-z)]X(z)$. Often $H_0(z)$ is a prototype low pass filter which implies that $H_1(z)$ is a high pass filter. Here let $H_0(z)$ be linear phase and FIR having the form

$$H_0(z) = \sum_{i=0}^n h_0(i)z^{-i}, \quad h_0(i) = h_0(n-i), \quad (2.2)$$

where $i = 0, 1, \dots, \frac{n+1}{2}$. Thus the Fourier transform of $\{h_0(i)\}$ has a linear phase: $[H_0(e^{j\omega})]^2 = |H_0(e^{j\omega})|^2 e^{-jn\omega}$. It follows that

$$\hat{X}(e^{j\omega}) = e^{-jn\omega} [|H_0(e^{j\omega})|^2 - (-1)^n |H_0(e^{j(\omega+\pi)})|^2] X(e^{j\omega}). \quad (2.3)$$

For the case of odd n , the perfect reconstruction requires that the transfer function

$$T(e^{j\omega}) := \begin{bmatrix} H_0(e^{-j\omega}) & H_0(e^{-j(\omega+\pi)}) \end{bmatrix} \begin{bmatrix} H_0(e^{j\omega}) \\ H_0(e^{j(\omega+\pi)}) \end{bmatrix} = 1, \quad \forall \omega. \quad (2.4)$$

Define E_r , E_s as the reconstruction error and the stop band error respectively, We design $H_0(z)$ such that it minimizes the error function [36]

$$E = (1 - \alpha)E_r + \alpha E_s, \quad 0 < \alpha < 1, \quad (2.5)$$

$$E_r = \frac{1}{\pi} \int_0^\pi |T(e^{j\omega}) - 1|^2 d\omega = \frac{1}{\pi} \int_0^\pi ([|H_0(e^{j\omega})|^2 + |H_0(e^{j(\omega+\pi)})|^2] - 1)^2 d\omega, \quad (2.6)$$

$$E_s = \frac{1}{\pi} \int_{\omega_s}^\pi |H_0(e^{j\omega})|^2 d\omega \quad (2.7)$$

with ω_s the frequency of the stop band edge.

In the rest of the section, explicit formulas are derived for the error functions E_r and E_s , and thus E as defined in (2.5). We begin with stop band error E_s in (2.7) by letting

$$|H_0(e^{j\omega})| = 2h^T C(\omega), \quad h = \begin{bmatrix} h_0(0) \\ h_0(1) \\ \vdots \\ h_0(m-1) \end{bmatrix}, \quad C(\omega) = \begin{bmatrix} \cos(m-1/2)\omega \\ \cos(m-3/2)\omega \\ \vdots \\ \cos(1-1/2)\omega \end{bmatrix}, \quad (2.8)$$

where $m = \frac{n+1}{2}$. Thus,

$$E_s = h^T Q_s h, \quad Q_s = [Q_s(i, k)]_{i,k=1}^{m,m}, \quad (2.9)$$

where Q_s is an $m \times m$ matrix, and $Q_s(i, k)$ is the element of Q_s at i th row and k th column. It can be easily verified that for $i, k = 1, 2, \dots, m$, $Q_s(i, k)$'s are given by

$$\begin{aligned} Q_s(i, k) &= \frac{4}{\pi} \int_{\omega_s}^{\pi} \left(\cos\left(m - \frac{2i-1}{2}\omega\right) \cos\left(m - \frac{2k-1}{2}\omega\right) \right) d\omega \\ &= \frac{2}{\pi} \int_{\omega_s}^{\pi} (\cos(2m - i - k + 1)\omega + \cos(k - i)\omega) d\omega \end{aligned} \quad (2.10)$$

such that

$$Q_s(i, k) = -\frac{2}{\pi} \left(\frac{\sin(k-i)\omega_s}{k-i} + \frac{\sin(2m-i-k+1)\omega_s}{2m-i-k+1} \right), \quad i \neq k, \quad (2.11)$$

$$Q_s(i, i) = \frac{2}{\pi} \left(\pi - \omega_s - \frac{\sin 2(m-i+0.5)\omega_s}{2(m-i+0.5)} \right). \quad (2.12)$$

The above expressions are essentially the same as those in [34]. The derivation for the reconstruction error E_r is more involved. Therefore, a system approach is adopted

to tackle the issue. Lyapunov equation is used to derive an explicit expression for E_r . For this purpose, we associate $H_0(z)$ with a realization

$$H_0(z) = D_0 + C_0(zI - A_0)^{-1}B_0$$

where A_0 has dimension $n \times n$ and both B_0 and C_0^T are vectors of size n given by

$$A_0 = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}, \quad B_0 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (2.13)$$

$$C_0 = \begin{bmatrix} h_0(1) & h_0(2) & \cdots & h_0(n) \end{bmatrix}, \quad D_0 = \begin{bmatrix} h_0(0) \end{bmatrix}. \quad (2.14)$$

It follows that $H_1(z) = H_0(-z) = D_0 - C_0(zI + A_0)^{-1}B_0$. Define a similarity transformation matrix $T = \text{diag}\{(-1)^{k-1}\}_{k=1}^n = \text{diag}\{1, -1, 1, \dots\}$. Then there hold $T^{-1}A_0T = -A_0$ and $T^{-1}B_0 = B_0$ which implies that

$$H_1(z) = D_0 - C_0T(zI - A_0)^{-1}B_0.$$

Now define

$$H(z) = \begin{bmatrix} H_0(z) \\ H_1(z) \end{bmatrix} = \begin{bmatrix} D_0 \\ D_0 \end{bmatrix} + \begin{bmatrix} C_0 \\ -C_0T \end{bmatrix} (zI - A_0)^{-1}B_0, \quad (2.15)$$

$$A = A_0, \quad B = B_0, \quad C = \begin{bmatrix} C_0 \\ -C_0T \end{bmatrix}, \quad D = \begin{bmatrix} D_0 \\ D_0 \end{bmatrix}.$$

For the perfect reconstruction case, $|H_0(z)|^2 + |H_1(z)|^2 = 1$ which implies that $H^T(z^{-1})H(z) = 1$. The reconstruction error can thus be written as

$$E_r = \frac{1}{2\pi} \int_0^{2\pi} |1 - H^T(e^{-j\omega})H(e^{j\omega})|^2 d\omega.$$

Because A is stable (all poles are at the origin), the following Lyapunov equation

$$X - A^T X A = C^T C \quad (2.16)$$

has a unique symmetric and positive definite solution [37]. Moreover, the simple realizations in (2.13) and (2.15) implies that $(zI - A)^{-1}B = [z^{n-1} \ \dots \ z \ 1]^T/z^n$. Then according to [24],

$$\begin{aligned} & B^T(z^{-1}I - A^T)^{-1}A^T X B + B^T X A(zI - A)^{-1}B + B^T X B \\ &= B^T(z^{-1}I - A^T)^{-1}C^T C(zI - A)^{-1}B, \end{aligned}$$

that in turn gives the expression

$$\begin{aligned} H^T(z^{-1})H(z) &= D^T D + B^T X B + (D^T C + B^T X A)(zI - A)^{-1}B \\ &\quad + B^T(z^{-1}I - A^T)^{-1}(C^T D + A^T X B), \\ 1 - H^T(z^{-1})H(z) &= v_0 + \sum_{k=1}^n v_k z^{-k} + \sum_{k=1}^n v_k z^k \end{aligned}$$

where v_0 and $v = [v_1 \ v_2 \ \dots \ v_n]^T$ are defined as

$$\begin{aligned} v_0 &= 1 - D^T D - B^T X B, \\ v &= C^T D + A^T X B. \end{aligned} \quad (2.17)$$

Applying the Parseval's Theorem, $E_r = |v_0|^2 + 2 \sum_{k=1}^n |v_k|^2$. In order to obtain an explicit expression, the solution of the Lyapunov equation $X - A^T X A = C^T C$ is required. It turns out that the elements of X has a nice form in terms of the prototype filter coefficient $\{h_0(i)\}$. Indeed denote $X = [x_{i,k}]$ with $x_{i,k} = x_{k,i}$ being the i th row and k th column element of X . Then $x_{n,k}$ can be obtained explicitly as $x_{n,k} = h_0(n)h_0(k) (1 + (-1)^{n+k})$, for $1 \leq k \leq n$ [22]. And the general $x_{i,k}$ can be obtained inductively as

$$\begin{aligned} x_{i,k} &= x_{n,n+k-i} + (1 + (-1)^{i+k}) \sum_{l=0}^{n-i-1} h_0(i+l)h_0(k+l) \\ &= (1 + (-1)^{i+k}) \sum_{l=0}^{n-i} h_0(i+l)h_0(k+l) \\ &= \begin{cases} 0 & \text{if } i+k = \text{odd}, \\ 2 \sum_{l=0}^{n-i} h_0(i+l)h_0(k+l) & \text{if } i+k = \text{even}. \end{cases} \end{aligned}$$

On the other hand,

$$B^T X B = x_{11} = 2 \sum_{l=0}^{n-1} h_0^2(1+l).$$

For the case considered in this chapter, n is odd. (In fact the expression derived for odd n also applies to the case of even n). Then

$$\begin{aligned} D^T C &= \begin{bmatrix} h_0(0) & h_0(0) \end{bmatrix} \begin{bmatrix} h_0(1) & h_0(2) & \cdots & h_0(n) \\ -h_0(1) & h_0(2) & \cdots & (-1)^n h_0(n) \end{bmatrix} \\ &= \begin{bmatrix} 0 & 2h_0(0)h_0(2) & 0 & 2h_0(0)h_0(4) & \cdots & 0 & 2h_0(0)h_0(n-1) & 0 \end{bmatrix}, \end{aligned}$$

$$B^T X A = \begin{bmatrix} x_{1,2} & x_{1,3} & \cdots & x_{1,n} & 0 \end{bmatrix} = \begin{bmatrix} 0 & x_{1,3} & 0 & x_{1,5} & \cdots & 0 & x_{1,n} & 0 \end{bmatrix}.$$

Now the vector v as in (2.17) are given by $v_{2k-1} = 0$ and $v_{2k} = 2h_0(0)h_0(2k) + x_{1,2k+1} = 2h_0(0)h_0(2k) + 2 \sum_{l=0}^{n-2k-1} h_0(1+l)h_0(2k+1+l) = 2 \sum_{l=0}^{n-2k} h_0(l)h_0(2k+l)$, where $k = 1, 2, \dots, m-1$ with $m = (n+1)/2$. Define a shifting matrix S of size $2m \times 2m$ and a vector h_0 of size $2m$ as

$$S = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \cdots & 0 & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix}, \quad h_0 = \begin{bmatrix} h_0(0) \\ h_0(1) \\ \vdots \\ h_0(2m-2) \\ h_0(2m-1) \end{bmatrix}. \quad (2.18)$$

Then it follows that

$$\begin{aligned} v^T v &= \|v\|_2^2 = \|D^T C + B^T X A\|_2^2 = \sum_{k=1}^{m-1} (v_{2k})^2 \\ &= \sum_{k=1}^{m-1} \left(2 \sum_{l=0}^{n-2k} h_0(l)h_0(2k+l) \right)^2 = 4 \sum_{k=1}^{m-1} (h_0^T S^{2k} h_0)^2. \end{aligned}$$

Regarding the constant term, by (2.13), we have

$$D^T D + B^T X B = 2h_0^2(0) + x_{1,1} = 2h_0^2(0) + 2 \sum_{l=0}^{n-1} h_0^2(1+l) = 2h_0^T h_0.$$

Thus the final expression for the reconstruction error is

$$\begin{aligned} E_r &= \|1 - (D^T D + B^T X B)\|_2^2 + 2\|D^T C + B^T X A\|_2^2 \\ &= (1 - 2h_0^T h_0)^2 + 2 \times 4 \sum_{k=1}^{m-1} (h_0^T S^{2k} h_0)^2. \end{aligned}$$

Summarizing the above derivation gives the following result for the reconstruction error.

Proposition 2.1 *Let S and h_0 be defined as in (2.18). Then the reconstruction error (2.6) is given by*

$$E_r = (1 - 2h_0^T h_0)^2 + 8 \sum_{i=1}^{m-1} \left(h_0^T S^{2i} h_0 \right)^2.$$

It is noted that the expression for E_r is essentially the same as the one in [70]. The drawback of the expression in [70] lies in its summation form of each individual coefficient which is extremely complex and thus difficult to be analyzed. On the contrary, the expression for E_r obtained in Proposition 2.1 has a much simpler form that allows us to propose new convergent iterative algorithms for the corresponding design problem. Let

$$h_0 = Mh, \quad M = \begin{bmatrix} I_m \\ \tilde{I}_m \end{bmatrix}, \quad \tilde{I}_m = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ \vdots & \cdots & 1 & 0 \\ \vdots & \cdots & \cdots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix}, \quad (2.19)$$

and I_m is the identity matrix of size m . Then the final expression for the error function is

$$\begin{aligned} E &= (1 - \alpha) \left[(1 - 2h^T M^T M h)^2 + 8 \sum_{i=1}^{m-1} \left(h^T M^T S^{2i} M h \right)^2 \right] + \alpha h^T Q_s h \\ &= (1 - \alpha) \left[(1 - 4h^T h)^2 + 8 \sum_{i=1}^{m-1} \left(h^T M^T S^{2i} M h \right)^2 \right] + \alpha h^T Q_s h, \end{aligned} \quad (2.20)$$

due to $M^T M = 2I_m$. The design problem is now converted to the synthesis of a vector h with size m such that E is minimized. Clearly this is a nonlinear optimization

problem. In the next section, we consider a more general nonlinear optimization problem and propose an iterative algorithm that is generalized from the one developed in [2]. The associated convergence issue will be studied accordingly.

2.3 AN ITERATIVE ALGORITHM AND ITS CONVERGENCE ANALYSIS

This section considers a nonlinear optimization problem of the form

$$J(x) = \sum_{k=0}^{m-1} (x^T P_k x)^2 + x^T Q x = x^T P[x, x] x + x^T Q x, \quad P[x, x] = \sum_{k=0}^{m-1} P_k x x^T P_k, \quad (2.21)$$

where x has dimension d , and $P_k = P_k^T$, $Q = Q_+ + Q_-$ with $Q_+ = Q_+^T$ positive definite, denoted by $Q_+ > 0$, and $Q_- = Q_-^T$ negative semidefinite, denoted by $Q_- \leq 0$. This index function is similar to the one in (2.20). In fact,

$$J(h) = E - (1 - \alpha)$$

by taking $d = m$ and

$$Q_+ = \alpha Q_s, \quad Q_- = -4(1 - \alpha) M^T M = -8(1 - \alpha) I_m, \quad (2.22)$$

$$P_0 = 2\sqrt{1 - \alpha} M^T M = 4\sqrt{1 - \alpha} I_m, \quad (2.23)$$

$$P_k = \sqrt{2(1 - \alpha)} M^T (S^{2k} + (S^T)^{2k}) M, \quad (2.24)$$

where $k = 1, 2, \dots, m - 1$ by $M^T M = 2I_m$. Thus minimization of E is a special case of minimization of $J(x)$. It should be clear that the iterative algorithm in [2] is not applicable to this more general nonlinear optimization problem in the form of (2.21). We propose the following new iterative algorithm generalized from [2].

Iterative Algorithm (A):

- Step 1: Choose initial guess $x = x_0 \neq 0$ such that $J_0 = \tilde{J}(x_0) < 0$.
- Step 2: For $i = 0, 1, 2, \dots$, do the following:
 - Step 2a: Find $x = x_{i+1}$ such that it minimizes the auxiliary index function

$$\tilde{J}_i[x_i](x) = x^T \left(P[x_i, x_i] + \frac{1}{2} Q_+ \right) x + x_i^T Q_- x + \frac{1}{2} x_i^T Q_+ x_i. \quad (2.25)$$

The optimal solution is given by

$$x = x_{i+1} = - (2P[x_i, x_i] + Q_+)^{-1} Q_- x_i. \quad (2.26)$$

- Step 2b: If $\|x_{i+1} - x_i\| \leq \epsilon$, with ϵ a pre-specified error tolerance, stop;
Otherwise set $i := i + 1$ and repeat Step 2a.

End.

The motivation of the proposed iterative algorithm (A) is as follows: suppose that the minimum solution x_* is known, then the iterative algorithm in [2] can be improved with the following iterative algorithm

$$\min_x \hat{J}_i[x_*](x) = \min_x \left(x^T P[x_*, x_*] x + \frac{1}{2} x^T Q_+ x + x_i^T Q_- x + \frac{1}{2} x_*^T Q_+ x_* \right). \quad (2.27)$$

Note that the Hessian of $\hat{J}_i[x_*](x)$ is exactly $2P[x_*, x_*] + Q_+$ that is positive definite for all x . It follows that the minimum solution for $\hat{J}_i[x_*](x)$ is given by $x = x_{i+1} = - (2P[x_*, x_*] + Q_+)^{-1} Q_- x_i$. Its difference from (2.25)-(2.26) lies in the nature that x_* is unknown. The best one can do is to replace $P[x_*, x_*]$ by $P[x_i, x_i]$ and $x_*^T Q_+ x_*$

by $x_i^T Q_+ x_i$. If Newton's algorithm is used to solve $\tilde{J}_i[x_i](x)$, its Hessian is exactly $2P[x_i, x_i] + Q_+$. Thus we obtain a Newton's type algorithm.

It is noted that the inversion in (2.26) is well defined by the fact that $P[x, x]$ is positive semidefinite and Q_+ is positive definite. Thus there exists a unique optimal solution satisfying

$$\tilde{J}[x_i](x_{i+1}) = \min_x \tilde{J}[x_i](x). \quad (2.28)$$

It is also noted that the auxiliary index function is very similar to $J(x)$. In fact, there holds

$$J(x_i) = \tilde{J}x_i.$$

An immediate problem to the iterative algorithm (A) is whether or not it converges. This is answered by the following result.

Theorem 2.2 *An optimal solution $x = x_*$ for $J(x)$ given in (2.21) satisfies the equality*

$$(2P[x_*, x_*] + Q) x_* = 0. \quad (2.29)$$

Moreover, if $J(x)$ is bounded below for all x , then algorithm (A) defines a descend algorithm for seeking the minimum of $J(x)$ given in (2.21). Furthermore, the sequence of solutions $\{x_i\}$ given by (2.26) converges to a limit x that satisfies the necessary condition for optimality in (2.29).

Proof: If $x = x_*$ is optimal for $J(x)$, then its gradient is zero at $x = x_*$. Thus (2.29) is verified easily by the fact $\frac{\partial J}{\partial x}(x_*) = 2(2P[x_*, x_*] + Q) x_* = 0$. It follows that the gradient at i th iteration is given by $\frac{\partial \tilde{J}}{\partial x}(x_i) = 2[(2P[x_i, x_i] + Q_+) x_i + Q_- x_i]$. This gives the relation $(2P[x_i, x_i] + Q_+)^{-1} Q_- x_i = \frac{1}{2} (2P[x_i, x_i] + Q_+)^{-1} \frac{\partial \tilde{J}}{\partial x}(x_i) - x_i$. Substitute it into (2.26) yields $x_{i+1} = x_i - \frac{1}{2} (2P[x_i, x_i] + Q_+)^{-1} \frac{\partial \tilde{J}}{\partial x}(x_i)$. This

is equivalent to $-(2P[x_i, x_i] + Q_+)(x_{i+1} - x_i) = \frac{1}{2} \frac{\partial J}{\partial x}(x_i)$, and thus there holds $-(x_{i+1} - x_i)^T (2P[x_i, x_i] + Q_+)(x_{i+1} - x_i) = \frac{1}{2} (x_{i+1} - x_i)^T \frac{\partial J}{\partial x}(x_i) < 0$ since $Q_+ > 0$ and $P[x_i, x_i] \geq 0$. Therefore, the algorithm (A) defines an iterative descend algorithm to solve the minimization problem for $J(x)$. To show its convergence, it is noted that by (2.28) x_{i+1} minimizes $\tilde{J}[x_i](x)$ and x_i minimizes $\tilde{J}[x_{i-1}](x)$ as defined in (2.25). That is,

$$\begin{aligned} \tilde{J}[x_i](x_{i+1}) &= \min_x \tilde{J}[x_i](x) \leq \tilde{J}[x_i](x_{i-1}) \\ &= x_{i-1}^T P[x_i, x_i] x_{i-1} + \frac{1}{2} x_{i-1}^T Q_+ x_{i-1} + x_i^T Q_- x_{i-1} + \frac{1}{2} x_i^T Q_+ x_i \\ &= x_i^T P[x_{i-1}, x_{i-1}] x_i + \frac{1}{2} x_i^T Q_+ x_i + x_{i-1}^T Q_- x_i + \frac{1}{2} x_{i-1}^T Q_+ x_{i-1} \\ &= \tilde{J}[x_{i-1}](x_i) = \min_x \tilde{J}[x_{i-1}](x). \end{aligned}$$

In the above derivation, the property $x_{i-1}^T P[x_i, x_i] x_{i-1} = x_i^T P[x_{i-1}, x_{i-1}] x_i$ is used which is easy to show in light of (2.21). It follows that $\tilde{J}[x_i](x_{i+1}) \leq \tilde{J}[x_{i-1}](x_i)$. Hence $\{\tilde{J}[x_i](x_{i+1})\}$ is a decreasing sequence. Because $\tilde{J}x$ is bounded below by the fact that $\lim_{\|x\| \rightarrow \infty} \tilde{J}x = +\infty$ due to

$$\tilde{J}x = x^T \left(P[x, x] + \frac{1}{2} Q_+ \right) x + x^T Q_- x + \frac{1}{2} x^T Q_+ x = J(x),$$

there exists a unique limit x for $\{x_i\}$. This limit clearly satisfies the necessary condition for optimality (2.29) according to (2.26). Finally note that the assumption $J(x_0) < 0$ excludes the possibility of $x = 0$ as the limit. ■

The auxiliary index function in (2.25) is crucial in the development of our main theorem. The next result indicates that the proposed iterative algorithm (A) is quasi-Newton.

Proposition 2.3 *For the index function in (2.21), its Hessian is given by*

$$H(x) = \frac{\partial^2 J}{\partial x^2}(x) = Q + 4P[x, x] + 2\tilde{P}[x, x], \quad \tilde{P}[x, x] = \sum_{k=1}^m (x^T P_k x) P_k. \quad (2.30)$$

Thus the iterative algorithm (A) is a quasi-Newton and it converges quadratically in the neighborhood of the local optimum.

Proof: The expression for the Hessian can be easily verified. By the proof for Theorem 2.2, the iterative algorithm has the form

$$x_{i+1} = x_i - [\tilde{H}(x_i)]^{-1} \frac{\partial J}{\partial x}(x_i),$$

where $\tilde{H} = 2P[x, x] + Q_+$ can be considered as the positive definite part of the Hessian. Thus the iterative algorithm (A) is a quasi-Newton algorithm that converges quadratically in the neighborhood of local optimum [54]. ■

Although the iterative algorithm (A) is proven to be convergent with quadratic convergence rate if the initial guess is close to the optimal solution, it does not guarantee the fast convergence if the initial guess is far away from the optimal solution. In fact, it is often the case that the solution sequence converges rather slowly with randomly chosen initial solution. A simple explanation is that at step i , x_{i+1} can be considered as x_i plus a correction term and the correction is often excessive due to the large value of $\|(2P[x_i, x_i] + Q_+)^{-1} \frac{\partial J}{\partial x}(x_i)\|$ when x_i is far away from the optimal solution. A possible remedy is to set $y_{i+1} = \tau x_{i+1} + (1 - \tau)y_i$, with $0 < \tau < 1$, as the sequence of solutions in the iterative algorithm, where x_{i+1} is the optimal solution for the index function $\tilde{J}[x_i](x)$. We also need to point out that Algorithm (A) is different from the algorithm proposed in [2]. In particular

the error function in [2] is given by $\tilde{E}(\alpha) = E_r + \alpha E_s$. The next result clarifies the difference.

Proposition 2.4 *If the iterative algorithm proposed in [2] converges, then the limit satisfies the necessary optimality condition of $E(2\alpha)$.*

Proof: It is not difficult to show that the algorithm at stage i seeks to minimize

$$\hat{J}_i(x) = x^T P[x_i, x_i]x + x^T Q_+ x + x^T Q_- x_i,$$

for some P_k 's and Q_- , by the fact that $Q_+ = \alpha Q_s$. This gives the iterative algorithm

$$x_{i+1} = -(2P[x_i, x_i] + 2Q_+)^{-1} Q_- x_i.$$

Thus if the algorithm in [2] is convergent, then its limit x satisfies

$$0 = 2P[x, x]x + 2Q_+ x + Q_- x = 2P[x, x]x + 2\alpha Q_s x + Q_- x. \quad (2.31)$$

On the other hand, the necessary optimality condition of $E(2\alpha)$ has the same form:

$$\frac{\partial E(2\alpha)}{\partial x} = 4P[x, x]x + 4\alpha Q_s x + 2Q_- x = 0 \quad (2.32)$$

with $x^T P[x, x]x + x^T Q_- x = E_r - 1$. Thus (2.31) and (2.32) are the same that concludes the proof. ■

The above proposition indicates that the algorithm proposed in [2], though works for $\tilde{E}(\alpha)$ by setting α suitably, does not work for E as a convex combination of E_r and E_s .

2.4 A DIFFERENT ITERATIVE ALGORITHM AND ITS CONVERGENCE ANALYSIS

This section considers a constrained nonlinear optimization problem where the index function $J(x)$ is the same as in (2.21) except for imposing a constraint

$$x^T \Pi x = 1, \quad \Pi = \Pi^T > 0. \quad (2.33)$$

This constrained problem was initially proposed in [34] for the design of QMF bank. With a hard constraint $4h^T h = 1$, the error index function to be minimized in [34] is

$$E = \alpha h^T Q_s h + 8 \sum_{i=1}^{m-1} \left(h^T M^T S^{2i} M h \right)^2$$

(note the difference between the above h and the vector w in [34]). Let

$$Q = \alpha Q_s, \quad P[x, x] = \sum_{k=1}^{m-1} P_k x x^T P_k, \quad P_k = \sqrt{2} M^T \left(S^{2k} + (S^T)^{2k} \right) M,$$

the error index in [34] can be converted into the form of (2.21). An iterative algorithm is proposed in [34] in which the solution h is computed as an eigenvector corresponding to a minimum eigenvalue of a certain matrix at each step. In this section we propose a new algorithm generalized from that in [34] and investigate again the associated convergence issue. The proposed algorithm seeks minimizing the index function $J(x)$ in (2.21) with constraint (2.33), and is summarized as follows:

Iterative Algorithm (B):

- Step 1: Choose initial guess $x = x_0 \neq 0$ such that $J_0 = \bar{J}(x_0) < 0$.
- Step 2: For $i = 0, 1, 2, \dots$, do the following:

- Step 2a: Find $x = x_{i+1}$ such that it minimizes the auxiliary index function

$$\bar{J}[x_i](x) = x^T \left(P[x_i, x_i] + \frac{1}{2}Q \right) x + \frac{1}{2}x_i^T Q x_i - \lambda_{i+1} (x^T \Pi x - 1). \quad (2.34)$$

- Step 2b: If $\|x_{i+1} - x_i\| \leq \epsilon \|x_i - x_{i-1}\|$, with ϵ a prespecified error tolerance, stop; Otherwise set $i := i + 1$ and repeat Step 2a.

End.

It should be clear that the parameter λ_{i+1} in (2.34) at the i th iteration is the Lagrange multiplier, and the optimal solution $x = x_{i+1}$ at Step 2a satisfies $x_{i+1}^T \Pi x_{i+1} = 1$ and

$$\left(\lambda_{\min} I - \Pi^{-1}(2P[x_i, x_i] + Q) \right) x_{i+1} = 0 \quad (2.35)$$

where λ_{\min} is the smallest eigenvalue of $\Pi^{-1}(2P[x_i, x_i] + Q)$ ¹. The constraint in (2.33) is simply a normalization as the size of eigenvector which can be rescaled. However, the algorithm (B) is different from that in [34]. In particular, it is unclear whether the iterative algorithm in [34] is convergent or not. The next result is concerned with the convergence issue associated with the iterative algorithm (B). Therefore, it is another major point in this chapter.

Theorem 2.5 *Suppose that $J(x)$ is the index function to be minimized given in (2.21) with $Q > 0$, subject to the constraint (2.33). Then the iterative algorithm (B) converges to a local minimum of $J(x)$ subject to the constraint (2.33).*

¹Note that all the eigenvalues of $\Pi^{-1}(2P[x_i, x_i] + Q)$ are positive real due to $P[x_i, x_i] \geq 0$ and $Q > 0$, $\Pi > 0$.

Proof: Since $x_k^T \Pi x_k = 1$ for each k , and $x = x_{i+1}$ minimizes $\bar{J}[x_i](x)$, it follows that

$$\begin{aligned} \bar{J}[x_i](x_{i+1}) &\leq x_{i-1}^T \left(P[x_i, x_i] + \frac{1}{2}Q \right) x_{i-1} + \frac{1}{2}x_i^T Q x_i \\ &= x_i^T \left(P[x_{i-1}, x_{i-1}] + \frac{1}{2}Q \right) x_i + \frac{1}{2}x_{i-1}^T Q x_{i-1} = \bar{J}[x_{i-1}](x_i). \end{aligned}$$

Thus $\{\bar{J}[x_i](x_{i+1})\}$ is again a decreasing sequence. Because $\bar{J}x = J(x)$ is bounded below, there exists a unique limit $x = x_*$ such that $\bar{J}[x_i](x_{i+1})$ converges to $\bar{J}x_* = J(x_*) = x_*^T (P[x_*, x_*] + Q) x_*$ and satisfies the constraint $x_*^T \Pi x_* = 1$. Thus the iterative algorithm (B) is convergent. Moreover the limit $x = x_*$ satisfies $(\lambda_{\min} I - \Pi^{-1}(2P[x_*, x_*] + Q)) x_* = 0$ in light of (2.35) which is exactly the necessary condition for $J(x)$ subject to the constrain $x^T \Pi x = 1$. Since $Q > 0$, and $P[x_*, x_*] \geq 0$, the Hessian is positive definite which proves that it is a local minimum. \blacksquare

Finally, the following result is similar to Proposition (2.4) so the proof is omitted. It clarifies the difference between the proposed iterative algorithm (B) and the one in [34].

Proposition 2.6 *Define the error index as*

$$\tilde{E}(\alpha) = \alpha h^T Q_s h + 8 \sum_{i=1}^{m-1} \left(h^T M^T S^{2i} M h \right)^2 = \alpha h^T Q_s h + h^T P[h, h] h$$

under the constraint $4h^T h = 1$ which is the same as in [34]. If the iterative algorithm proposed in [34] converges, then the limit satisfies the necessary optimality condition of $E(2\alpha)$ under the constraint $4h^T h = 1$.

CHAPTER 3

TWO-CHANNEL QMF BANK DESIGN

3.1 SUMMARY

A popular approach to design QMF banks is the use of polyphase decomposition, described in [90], where both analysis and synthesis filter banks are designed in terms of their polyphase components. This is referred to as *indirect* approach. We adopt a direct approach to the design of QMF banks without using polyphase decomposition. Although it has several disadvantages as pointed out in [90], the necessary and sufficient condition for perfect reconstruction (PR) can be easily derived and used for the design of QMF filter banks. Moreover these disadvantages are insignificant if the analysis filter bank and synthesis filter bank are designed separately. Indeed in practice, analysis filter banks are often designed to satisfy frequency domain specifications required for subband coding. Additional requirement for perfect reconstruction may not be suitable in some applications. It is the goal of synthesis filter banks to satisfy the PR condition. If PR is not possible, then the synthesis filter bank should be designed such that it minimizes certain measure for the reconstruction error that is referred to as *model matching* [74, 3].

We consider the QMF bank as shown in Figure 2.1. Such multirate systems play an important role in subband coding of speech and digital audio application, and have been studied extensively; See, for example, [3, 13, 17, 47, 52, 62, 78, 90] and their corresponding references. Ideally $\hat{x}(n)$ is a delayed version of $x(n)$ that is defined as PR. However in practice PR may not be possible that will be the case when analysis filter bank is designed independent of the synthesis filter bank. We

consider the design issue for the synthesis filter bank such that certain measure of the error signal $e(n) = \hat{x}(n) - x(n - d)$ is minimized for some $d \geq 0$. Two cases will be studied. The first one is when the signal $x(n)$ is white noise and the goal is to minimize the covariance of the error that leads to \mathcal{H}_2 optimization. The second one is when $x(n)$ is energy bounded signal and the goal is to minimize the energy of the error in worst case that leads to \mathcal{H}_∞ optimization. Both are equivalent to optimizations in frequency domain. Optimal solutions will be derived and numerical algorithms for computing the optimal synthesis filter banks will be developed. Compared with the results in [3] where suboptimal \mathcal{H}_∞ solutions are used for polyphase components of the QMF bank, our design algorithm is more efficient numerically and it produces the optimal synthesis filter bank with simpler solutions and lower order. Furthermore, our direct approach removes the aliasing error completely while the indirect approach in [3] does not.

3.2 DESIGN OF SYNTHESIS FILTER BANKS WITH \mathcal{H}_2 AND \mathcal{H}_∞ OPTIMIZATIONS

In the existing literature, analysis and synthesis filter banks are often designed simultaneously. A new development is the work of [74, 3] where analysis and synthesis filter banks are designed separately. The underlying philosophy is that the analysis filter bank is designed to satisfy the specification in frequency domain required for subband coding, while the synthesis filter bank is designed to satisfy the PR condition. If PR is not possible, then the synthesis filter bank should be designed to minimize the model matching error. We are motivated by the work in [74, 3] but adopt a direct approach to the design of QMF bank that establishes stronger results than those in [74, 3].

3.2.1 FORMULATION OF THE DESIGN APPROACH

Consider the QMF bank in Figure 2.1 where all the filters are restricted to be both causal and stable. The input/output relation in \mathcal{Z} -transform is given by

$$\begin{aligned}\hat{X}(z) = & \frac{1}{2} [H_0(z)F_0(z) + H_1(z)F_1(z)] X(z) \\ & + \frac{1}{2} [H_0(-z)F_0(z) + H_1(-z)F_1(z)] X(-z).\end{aligned}\quad (3.1)$$

The term involving $X(-z)$ is called aliasing error. If the synthesis filters $F_0(z)$ and $F_1(z)$ can be chosen as $G(z)H_1(-z)$ and $-G(z)H_0(-z)$ respectively, with $G(z)$ a causal and stable transfer function, then the aliasing error will be eliminated completely and

$$\hat{X}(z) = T_r(z)X(z), \quad T_r(z) = \frac{1}{2} [H_0(z)F_0(z) + H_1(z)F_1(z)]. \quad (3.2)$$

In addition, if $T_r(z) = cz^{-d}$ for some $c \neq 0$ and $d \geq 0$, then $\hat{X}(z)$ is the perfect reconstruction of $X(z)$. The next theorem follows easily from this definition, and thus the proof is skipped.

Theorem 3.1 *The QMF bank shown in Figure 2.1 has perfect reconstruction property if and only if*

$$\begin{bmatrix} H_0(z) & H_1(z) \\ H_0(-z) & H_1(-z) \end{bmatrix} \begin{bmatrix} F_0(z) \\ F_1(z) \end{bmatrix} = \begin{bmatrix} 2cz^{-d} \\ 0 \end{bmatrix} \quad (3.3)$$

for some $c \neq 0$ and $d \geq 0$.

For the QMF bank in Figure 2.1, define the alias component (AC) matrix and synthesis matrix as

$$H_{AC}(z) = \begin{bmatrix} H(z) \\ H(-z) \end{bmatrix} = \begin{bmatrix} H_0(z) & H_1(z) \\ H_0(-z) & H_1(-z) \end{bmatrix}, \quad F(z) = \begin{bmatrix} F_0(z) \\ F_1(z) \end{bmatrix}. \quad (3.4)$$

Also define the analysis matrix $H(z)$ as the first row of $H_{AC}(z)$. If $\det H_{AC}(z) \neq 0$ for all $|z| \geq 1$, then the ideal synthesis matrix

$$F_{id}(z) := \begin{bmatrix} H_0(z) & H_1(z) \\ H_0(-z) & H_1(-z) \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\det H_{AC}(z)} \begin{bmatrix} H_1(-z) \\ -H_0(-z) \end{bmatrix} \quad (3.5)$$

is causal, stable, and achieves perfect reconstruction with $c = 0.5$ and $d = 0$. Readers may easily find that for $\det H_{AC}(z) = 1$, the above equation reduces to $F_0(z) = H_1(-z)$ and $F_1(z) = -H_0(-z)$ which is the conventional choice to cancel aliasing error. However, if $\det H_{AC}(z)$ has unstable zeros, then the ideal synthesis matrix $F_{id}(z)$ as in (3.5) is either noncausal or unstable. Because noncausal and unstable filters can not be used for real time implementation, PR is not possible for the two-channel QMF bank. In this case we design a causal and stable $G(z)$ to seek

$$F(z) = \begin{bmatrix} F_0(z) \\ F_1(z) \end{bmatrix} = G(z)(\det H_{AC}(z))F_{id}(z) = G(z) \begin{bmatrix} H_1(-z) \\ -H_0(-z) \end{bmatrix} \quad (3.6)$$

to minimize the covariance or the energy of the error signal $e(n)$ shown in Figure 3.1.

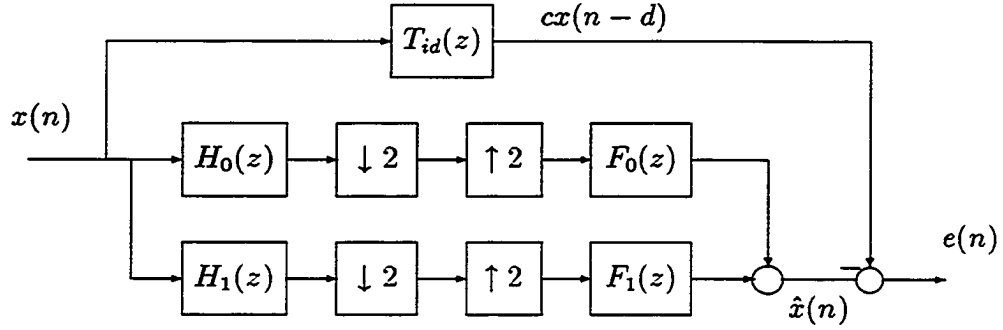


Figure 3.1: Two-channel error system.

Theorem 3.2 *Suppose that the analysis filter bank $H(z)$ is causal and stable for the QMF bank shown in Figure 3.1. Then all causal and stable synthesis filter banks $F(z)$ achieving zero aliasing error are parameterized as (3.6) for some causal and stable $G(z)$ that is real valued for each real z .*

Proof: Any causal and stable synthesis filter bank $F(z)$ achieving zero aliasing error satisfies $[H_0(-z) \ H_1(-z)] F(z) = 0$. Since $H(z)$ is causal and stable, $H(z)F(z) = \phi(z)$ is causal and stable as well. Combining this fact with (3.3) yields $F(z) = H_{AC}^{-1}(z) [1 \ 0]^T \phi(z) = F_{id}(z)\phi(z)$ with $F_{id}(z)$ as in (3.5). From the fact that $H(z)F_{id}(z) = 1$, and $H(z)$ is stable and causal, $F_{id}(z)$ has no unstable block zero. It follows from the causality and stability of $F(z)$ that $\phi(z) = G(z) \det H_{AC}(z)$ for some $G(z)$ that is also stable and causal. That is, the unstable poles of $F_{id}(z)$ that are also unstable zeros of $\det H_{AC}(z)$ have to be cancelled by unstable zeros of $\phi(z)$, in order to ensure $F(z)$ to be causal and stable. Substitute the expression of $\phi(z)$ into the expression of $F(z)$, $F(z) = F_{id}(z)G(z) \det H_{AC}(z) = G(z) [H_1(-z) \ -H_0(-z)]^T$ which is the same as (3.6). For real valued $H(z)$, $F(z)$ is real valued for each real value z if and only if $G(z)$ is. ■

Define $E(z) = \mathcal{Z}(e(n))$, the \mathcal{Z} -transform of the error signal $e(n)$ as in Figure 3.1. Then substitution of $F(z)$ into equation (3.1) gives

$$\begin{aligned} E(z) &= \left(\frac{1}{2} [H_0(z)H_1(-z) - H_1(z)H_0(-z)] G(z) - cz^{-d} \right) X(z) \\ &= T_E(z)X(z), \quad T_E(z) = \left(\frac{1}{2} G(z) \det H_{AC}(z) - cz^{-d} \right), \end{aligned} \quad (3.7)$$

by the fact that $\det H_{AC}(z) = H_0(z)H_1(-z) - H_1(z)H_0(-z)$ and the aliasing error is completely cancelled. Motivated by the model matching method in [74, 3], we seek a causal and stable transfer function $G(z)$ such that

$$\|T_E\|_2 := \sqrt{\frac{1}{2\pi} \int_0^{2\pi} |T_E(e^{j\omega})|^2 d\omega}, \quad \text{or} \quad \|T_E\|_\infty := \sup_{\omega} |T_E(e^{j\omega})|, \quad (3.8)$$

is minimized which is termed as \mathcal{H}_2 or \mathcal{H}_∞ optimization, respectively. Such optimizations are well known in the control community [104]. The synthesis filter bank can then be obtained according to (3.6). The symbol \mathcal{H}_2 or \mathcal{H}_∞ denotes the collection of all causal stable transfer functions $T(z)$ such that $\|T\|_2$ or $\|T\|_\infty$ is bounded, respectively.

Define auto-correlation and spectral density of a discrete-time signal $s(n)$ by $r_s(n) = \sum_{\tau=-\infty}^{\infty} s(\tau)s(\tau+n)$ and $R_s(\omega) = \sum_{n=-\infty}^{\infty} r_s(n)e^{-j\omega n}$. Denote spectral density of $x(n)$ by $R_x(\omega)$ and of $e(n)$ by $R_e(\omega)$. Then from [71] there hold $R_e(\omega) = |T_E(\omega)|^2 R_x(\omega)$ and $r_e(0) = \|e\|_2^2 = \sum_{n=-\infty}^{\infty} |e(n)|^2 = \frac{1}{2\pi} \int_0^{2\pi} R_e(\omega) d\omega$, where $r_e(0)$ is the covariance of the error function. If $x(n)$ is treated as stochastic signal that is assumed to be white noise with unit spectral density, then $r_x(n) = \delta(n)$ and $R_x(\omega) \equiv 1$. It follows that $\|e\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} R_e(\omega) d\omega = \frac{1}{2\pi} \int_0^{2\pi} |T_E(e^{j\omega})|^2 d\omega = \|T_E\|_2^2$, and minimization of the covariance of the error signal is equivalent to \mathcal{H}_2 optimization. If the design objective is to minimize the sup-norm (or Chebyshev norm) of the

spectral density of the error signal, then it results in \mathcal{H}_∞ optimization. There is also a deterministic interpretation for \mathcal{H}_∞ optimization. Suppose that the signal $x(n)$ has bounded energy and satisfies $\|x\|_2 = \sqrt{\sum_{n=-\infty}^{\infty} |x(n)|^2} < \infty$. Then the \mathcal{H}_∞ norm of $T_E(z)$ can be expressed as [104]

$$\|T_E\|_\infty = \sup_{\|x\|_2 \neq 0} \frac{\|e\|_2}{\|x\|_2}, \quad \|e\|_2 = \sqrt{\sum_{n=-\infty}^{\infty} |e(n)|^2}.$$

That is, the \mathcal{H}_∞ norm of $T_E(z)$ is the amplification of the energy from input signal $x(n)$ to output error $e(n)$ in the worst-case.

3.2.2 OPTIMAL \mathcal{H}_2 AND \mathcal{H}_∞ SOLUTIONS FOR THE SYNTHESIS FILTER BANK

Consider the transfer function $T_E(z)$ as in (3.7) and (3.8) where $\det H_{AC}(z)$ has exactly κ zeros outside of the unit circle, counting multiplicity. Then there exists a factorization

$$\det H_{AC}(z) = P(z)Q(z), \quad Q(z) = z^{-q} \frac{q_0 + q_1 z + \cdots + q_\kappa z^\kappa}{q_0 z^\kappa + q_1 z^{\kappa-1} + \cdots + q_\kappa z}, \quad q \geq 0, \quad (3.9)$$

where all the zeros of $Q(z)$ are outside of the unit circle and both $P(z)$ and $1/P(z)$ are causal and stable. Such a factorization is called inner-outer factorization with $Q(z)$ inner and $P(z)$ outer [104]. It is noted that $Q(z)Q(z^{-1}) \equiv 1$ by the allpass property, and thus

$$\|T_E\|_\alpha = \frac{1}{2} \|GPQ - 2T_{id}\|_\alpha = \frac{1}{2} \|GP - 2T_{id}Q^\sim\|_\alpha, \quad \alpha = 2, \infty, \quad (3.10)$$

where $Q^\sim = Q(z^{-1})$, and $T_{id}(z) = cz^{-d}$ is the ideal transfer function for PR.

In what follows next, we define projection operators Π_a and Π_c , and Hankel norm that will be used to obtain our main result. For an arbitrary transfer function $R(z)$ that is absolutely square-integrable, there exists a power series representation

$$R(z) = \sum_{k=-\infty}^{\infty} r(k)z^{-k}, \quad \|r\|_2^2 = \sum_{k=-\infty}^{\infty} |r(k)|^2 < \infty.$$

The projection operators Π_a and Π_c are defined by

$$\Pi_a[R] = R_a(z) = \sum_{k=1}^{\infty} r(-k)z^k, \quad \Pi_c[R] = R_c(z) = \sum_{k=0}^{\infty} r(k)z^{-k},$$

where $R_c(z)$ and $R_a(z)$ are the causal and anticausal parts of $R(z)$ respectively. For any transfer function $T(z)$ that has continuous frequency response, its Hankel norm is defined by [21, 104]:

$$\|T\|_H := \sup_{U(z) \in \mathcal{H}_2} \frac{\|\Pi_a[TU]\|_2}{\|U\|_2}. \quad (3.11)$$

Theorem 3.3 *Let $\det H_{AC}(z) = P(z)Q(z)$ be inner-outer factorization as in (3.9) where $Q(z)$ includes all the unstable zeros of $\det H_{AC}(z)$. Then the \mathcal{H}_2 optimal solution is given by*

$$G(z) = G_2(z) = \Pi_c[2T_{id}(z)Q^{\sim}(z)]/P(z), \quad \inf_{G(z) \in \mathcal{H}_2} \|T_E\|_2 = \|\Pi_a[T_{id}Q^{\sim}]\|_2, \quad (3.12)$$

with $T_{id}(z) = cz^{-d}$, and the \mathcal{H}_{∞} optimal solution by

$$\begin{aligned} G(z) = G_{\infty}(z) &= \Pi_c[2T_{id}(z)Q^{\sim}(z)]/P(z) \\ &\quad + \operatorname{argmin} \left\{ G_o(z) : \inf_{G_o \in \mathcal{H}_{\infty}} \|G_o - \Pi_a[2T_{id}Q^{\sim}]\|_{\infty} \right\} / P(z), \\ \inf_{G \in \mathcal{H}_{\infty}} \|T_E\|_{\infty} &= \|\Pi_a[T_{id}Q^{\sim}]\|_H. \end{aligned} \quad (3.13)$$

Proof: By (3.10), $2\|T_E\|_\alpha = \|\tilde{T}_E\|_\alpha$, and $\tilde{T}_E(z) = G(z)P(z) - 2T_{id}(z)Q^\sim(z) = \tilde{G}(z) - 2T_{id}(z)Q^\sim(z)$, where $T_{id}(z) = cz^{-d}$, $\alpha = 2$ or ∞ , and $\tilde{G}(z) = G(z)P(z)$. Since both $P(z)$ and $1/P(z)$ are causal and stable, $G(z)$ is causal and stable if and only if $\tilde{G}(z)$ is. Thus minimization of $\|T_E\|_\alpha$ is equivalent to minimization of $\|\tilde{T}_E\|_\alpha = \|\tilde{G} - 2T_{id}Q^\sim\|_\alpha$ subject to the condition that $\tilde{G}(z)$ is causal and stable. For the case $\alpha = 2$,

$$\begin{aligned}\tilde{T}_E(z) &= \tilde{G}(z) - 2T_{id}(z)Q^\sim(z) \\ &= (\tilde{G}(z) - \Pi_c[2T_{id}(z)Q^\sim(z)]) - (\Pi_a[2T_{id}(z)Q^\sim(z)]). \quad (3.14)\end{aligned}$$

By the orthogonality of \mathcal{H}_2 and \mathcal{H}_2^\perp , the Hilbert space perpendicular to \mathcal{H}_2 , there holds $\|\tilde{T}_E\|_2^2 = \|\tilde{G} - \Pi_c[2T_{id}Q^\sim]\|_2^2 + \|\Pi_a[2T_{id}Q^\sim]\|_2^2$. Hence the optimal \mathcal{H}_2 solution is given by

$$G(z) = G_2(z) = \tilde{G}_2(z)/P(z), \quad \tilde{G}_2(z) = \Pi_c[2T_{id}(z)Q^\sim(z)],$$

which gives the expression in (3.12). For the case $\alpha = \infty$, we set

$$\tilde{G}(z) = G_o(z) + \Pi_c[2T_{id}(z)Q^\sim(z)]$$

and thus $\tilde{T}_E(z)$ as in (3.14) can be rewritten as

$$\tilde{T}_E(z) = \tilde{G}(z) - 2T_{id}(z)Q^\sim(z) = G_o(z) - \Pi_a[2T_{id}(z)Q^\sim(z)],$$

where $\Pi_a[2T_{id}(z)Q^\sim(z)]$ is anti-causal but stable, and $\tilde{G}(z)$ is both causal and stable. It follows from Nehari's Theorem [104] that the infimum of $\|\tilde{T}_E\|_\infty$ over all causal

and stable $G_o(z)$ is the optimal Hankel norm approximation of $\Pi_a [2T_{id}(z)Q^\sim(z)]$ which gives the expression in (3.13). ■

We would like to point out that Theorem 3.3 applies to the indirect approach in [3] as well with minor modifications although the optimal \mathcal{H}_∞ optimal solution is not unique. The difference between our main result and that in [3] is that optimal solutions are derived explicitly in Theorem 3.3 and it can be used to compute the optimal synthesis filter bank which will be shown in the next section, while [3] uses the μ -toolbox in MATLAB to compute a suboptimal solution. The direct approach employed in this chapter also simplifies the optimal solutions significantly. The next result indicates that the optimal solutions yield PR asymptotically as d goes to infinity which is parallel to that of [3].

Theorem 3.4 *Suppose that $\det H_{AC}(z)$ avoids zeros on the unit circle. Then with the optimal \mathcal{H}_2 and \mathcal{H}_∞ solutions in Theorem 3.3, the reconstruction error approaches to zero as $d \rightarrow \infty$.*

Proof: Since $\det H_{AC}(z)$ has no zeros on unit circle, $Q^\sim(z)$ has all k poles strictly outside of unit circle and thus is analytic for all $|z| \leq 1$. It follows that there admits expansion

$$2T_{id}(z)Q^\sim(z) = 2cz^{-d} \sum_{k=0}^{\infty} \tilde{q}(k)z^k, \quad \sum_{k=0}^{\infty} |\tilde{q}(k)|^p < \infty,$$

for any $p \geq 1$. Its anticausal part is given by

$$2\Pi_a [T_{id}(z)Q^\sim(z)] = 2c \sum_{k=1}^{\infty} \tilde{q}(k+d)z^k \rightarrow 0$$

as $d \rightarrow \infty$ by the summability of $\tilde{q}(t)$. Thus the reconstruction error $e(t)$ approaches to zero asymptotically as $d \rightarrow \infty$. ■

The assumption of $\det H_{AC}(z)$ avoiding zeros on the unit circle is mild as explained in [3]. We will not elaborate further here.

3.2.3 DESIGN ALGORITHMS FOR THE SYNTHESIS FILTER BANK

The results in the previous two subsections give two design algorithms for synthesis filter banks based on \mathcal{H}_2 and \mathcal{H}_∞ optimizations. We again assume that $T_{id}(z) = cz^{-d}$ where $c > 0$ and d is to be determined in the design process. The first algorithm is based on \mathcal{H}_2 optimization.

Design Algorithm 1 (\mathcal{H}_2 Optimization):

- Step 1: For the given analysis filters $H_0(z)$ and $H_1(z)$ and design specification ϵ_2 , compute $\det H_{AC}(z) = H_0(z)H_1(-z) - H_1(z)H_0(-z)$.
- Step 2: Compute inner-outer factorization $\det H_{AC}(z) = P(z)Q(z)$ where $Q(z)$ is allpass function of the form in (3.9) that contains all unstable zeros of $\det H_{AC}(z)$.
- Step 3: Compute the optimal \mathcal{H}_2 performance index: $2\|\Pi_a[T_{id}Q^\sim]\|_2$.
- Step 4: If $2\|\Pi_a[T_{id}Q^\sim]\|_2 > \epsilon_2$, then increase d and go to Step 3; Otherwise go to Step 5.
- Step 5: Compute $G(z) = G_2(z) = \Pi_c[2T_{id}(z)Q^\sim(z)]/P(z)$ and set synthesis filter bank according to (3.6).

The second design algorithm is based on \mathcal{H}_∞ optimization.

Design Algorithm 2 (\mathcal{H}_∞ Optimization):

- Step 1: For the given analysis filters $H_0(z)$ and $H_1(z)$ and design specification ϵ_∞ , compute $\det H_{AC}(z) = H_0(z)H_1(-z) - H_1(z)H_0(-z)$.

- Step 2: Compute inner-outer factorization $\det H_{AC}(z) = P(z)Q(z)$ where $Q(z)$ is allpass function of the form in (3.9) that contains all unstable zeros of $\det H_{AC}(z)$.
- Step 3: Compute the optimal \mathcal{H}_∞ performance index: $2\|\Pi_a [T_{id}Q^\sim]\|_H$.
- Step 4: If $2\|\Pi_a [T_{id}Q^\sim]\|_H > \epsilon_\infty$, then increase d and go to Step 3; Otherwise go to Step 5.
- Step 5: Compute $G(z) = G_\infty(z)$ according to (3.13) and set synthesis filter bank according to (3.6).

Although the two algorithms seem straightforward, their numerical implementation with good numerical property and efficiency is not that simple. In the next section we will study the computational issue involved in each step of the two design algorithms in detail.

3.3 STATE-SPACE COMPUTATION

\mathcal{H}_2 and \mathcal{H}_∞ optimizations are developed in the control community using state-space computation tool. Although both of the proposed algorithms can be implemented without state-space computation, state-space method does improve numerical property of the computation if realizations are suitably chosen.

3.3.1 STATE-SPACE REALIZATIONS

We begin with state-space realization for the analysis matrix $H(z)$. Suppose $H(z)$ is of the fractional form $H(z) = D_h + M_h^{-1}(z)N_h(z)$, $M_h(z) = z^n + \sum_{i=1}^n m_i z^{n-i}$, and $N_h(z) = \sum_{i=1}^n n_i z^{n-i}$, where D_h and n_i vector of size 1×2 , $i = 1, \dots, n$. If both $H_0(z)$ and $H_1(z)$ are FIR filter, then $m_i = 0$ for $1 \leq i \leq n$. A simple realization of $H(z)$

can be obtained according to [37] in which a so-called observer form is adopted. The corresponding matrices are listed as

$$A_h = \begin{bmatrix} -m_1 & 1 & 0 & \cdots & 0 \\ -m_2 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 1 \\ -m_n & 0 & \cdots & \cdots & 0 \end{bmatrix}, \quad B_h = \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_n \end{bmatrix}, \quad C_h = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \end{bmatrix}. \quad (3.15)$$

Hence $H(z) = D_h + C_h(zI - A_h)^{-1}B_h$, and the realization (C_h, A_h) is observable.

With the above realization of $H(z)$ we can then obtain a realization for

$$\tilde{H}(z) = \begin{bmatrix} H_1(-z) \\ -H_0(-z) \end{bmatrix} = J \left(D_h^* - B_h^*(zI + A_h^*)^{-1}C_h^* \right), \quad J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

where “*” denotes conjugate and transpose. Thus a realization of $\det H_{AC}(z)$ is obtained as [104, 3]:

$$\det H_{AC}(z) = H(z)\tilde{H}(z) = D_{AC} + C_{AC}(zI - A_{AC})^{-1}B_{AC}, \quad D_{AC} = D_h J D_h^* = 0,$$

$$A_{AC} = \begin{bmatrix} A_h & -B_h J B_h^* \\ 0 & -A_h^* \end{bmatrix}, \quad B_{AC} = \begin{bmatrix} B_h J D_h^* \\ C_h^* \end{bmatrix}, \quad C_{AC} = \begin{bmatrix} C_h & -D_h J B_h^* \end{bmatrix}.$$

If minimal realization is preferred, balanced truncation [104] can be employed to eliminate the uncontrollable and unobservable modes of $\det H_{AC}(z)$. Actually by doing so, the computational effort will be reduced. The only thing need to be considered is the accuracy of the final results.

3.3.2 INNER-OUTER FACTORIZATION

Since $\det H_{AC}(z)$ is a scalar function and is strictly proper, there exists an integer $q > 0$ such that $C_{AC} A_{AC}^{q-1} B_{AC} \neq 0$ while $C_{AC} A_{AC}^i B_{AC} = 0$ for $0 \leq i \leq q-2$. Thus there holds

$$\begin{aligned} \tilde{T}(z) &= z^q \det H_{AC}(z) = \tilde{D} + \tilde{C}(zI - \tilde{A})^{-1} \tilde{B}, \\ \tilde{A} &= A_{AC}, \quad \tilde{B} = B_{AC}, \quad \tilde{C} = C_{AC} A_{AC}^q, \quad \tilde{D} = C_{AC} A_{AC}^{q-1} B_{AC} \neq 0. \end{aligned} \quad (3.16)$$

For $\tilde{T}(z) = z^q \det H_{AC}(z)$ in (3.16), its inverse with the corresponding realization are given by

$$\begin{aligned} T(z) &= 1/\tilde{T}(z) = 1 / (z^q \det H_{AC}(z)) = D + C(zI - A)^{-1} B, \\ A &= \tilde{A} - \tilde{B} \tilde{D}^{-1} \tilde{C}, \quad B = \tilde{B} \tilde{D}^{-1}, \quad C = -\tilde{D}^{-1} \tilde{C}, \quad D = \tilde{D}^{-1}. \end{aligned} \quad (3.17)$$

Since $1/T(z)$ is stable, the unstable modes of $T(z)$ are both controllable and observable which implies that (A, B) is stabilizable and (C, A) is detectable [37]. Hence there exists a unique stabilizing solution $X \geq 0$ to the ARE (algebraic Riccati equation)

$$A^*(I + XBB^*)^{-1}XA - X = 0. \quad (3.18)$$

An inner-outer factorization of $T(z)$ is now given by, using the formulas in [104],

$$\begin{aligned} T(z) &= N(z)M^{-1}(z), \\ \begin{bmatrix} M(z) \\ N(z) \end{bmatrix} &= \begin{bmatrix} R^{-1/2} \\ DR^{-1/2} \end{bmatrix} + \begin{bmatrix} K \\ C + DK \end{bmatrix} (zI - A - BK)^{-1} BR^{-1/2}, \end{aligned} \quad (3.19)$$

where $M(z)$ is an inner, or allpass factor, and $N(z)$ outer, $R = I + B^*XB$, and $K = -R^{-1}B^*XA = -(I + B^*XB)^{-1}B^*XA$. The inner-outer factorization in (3.9) can then be obtained as

$$\det H_{AC}(z) = P(z)Q(z), \quad P(z) = 1/N(z), \quad Q(z) = z^{-q}M(z).$$

The problem with the procedure from [104] is that the inner, or allpass factor $Q(z)$ has the same order as $\det H_{AC}(z)$, instead of κ . In the rest of the subsection, we will seek a minimal realization of $Q(z)$.

By Schur decomposition, there exists a real unitary matrix U such that $UU^* = U^*U = I$ and

$$\bar{A} = UAU^* = \begin{bmatrix} A_s & A_{su} \\ 0 & A_u \end{bmatrix}, \quad \bar{B} = UB = \begin{bmatrix} B_s \\ B_u \end{bmatrix}, \quad \bar{C} = CU = \begin{bmatrix} C_s & C_u \end{bmatrix} \quad (3.20)$$

where A_u of size $\kappa \times \kappa$ is completely unstable. In this case the ARE in (3.18) can be written as

$$\bar{A}^*(I + \bar{X}\bar{B}\bar{B}^*)^{-1}\bar{X}\bar{A} - \bar{X} = 0. \quad (3.21)$$

Proposition 3.5 *The stabilizing solution \bar{X} of the algebraic riccati equation (ARE) in (3.21) is given by*

$$\bar{X} = \begin{bmatrix} 0 & 0 \\ 0 & Z^{-1} \end{bmatrix}, \quad Z = A_u Z A_u^* - B_u B_u^*. \quad (3.22)$$

Proof: The Lyapunov equation in (3.22) is equivalent to $Z = A_u^{-1}Z(A_u^{-1})^* - A_u^{-1}B_u(A_u^{-1}B_u)^*$. Since A_u is completely unstable, A_u^{-1} is stable. The stabilizability

of (A, B) implies that (A_u, B_u) is controllable, and that in turn implies that the solution $Z > 0$. Hence \bar{X} exists and is positive semidefinite. Substituting \bar{X} in (3.22) into ARE (3.21) gives $A_u^*(I + Z^{-1}B_uB_u^*)^{-1}Z^{-1}A_u - Z^{-1} = 0$, which is the same as the Lyapunov equation in (3.22). Thus \bar{X} in (3.22) is a solution to (3.21). Since the above ARE can be written as $A_c^*Z^{-1}A_c - Z^{-1} + A_c^*Z^{-1}B_uB_u^*Z^{-1}A_c = 0$ with $A_c = (I + B_uB_u^*Z^{-1})^{-1}A_u$, an application of Lyapunov stability theorem concludes the stability of A_c . It follows that

$$A + BK = U^*(\bar{A} + \bar{B}\bar{K})U, \quad \bar{A} + \bar{B}\bar{K} = (I + \bar{B}\bar{B}^*\bar{X})^{-1}\bar{A} = \begin{bmatrix} A_s & A_{sc} \\ 0 & A_c \end{bmatrix}$$

for some A_{sc} . Thus \bar{X} in (3.22) is indeed a stabilizing solution to ARE (3.20) which is unique [104]. ■

Proposition 3.5 demonstrates that the ARE (3.18) can be converted into Lyapunov equation of lower order. With $Z > 0$ solved from (3.22), the stabilizing solution X to ARE (3.18) is given by

$$X = U_2^*Z^{-1}U_2, \quad U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}, \quad (3.23)$$

with U_2 of size $k \times n$, and the feedback gain \bar{K} is given by

$$\bar{K} = KU^* = \begin{bmatrix} 0 & K_u \end{bmatrix}, \quad K_u = -(1 + B_u^*Z^{-1}B_u)^{-1}B_u^*Z^{-1}A_u. \quad (3.24)$$

Simple calculation gives the minimal realization for the inner factor:

$$M(z) = (I + K_u(zI - A_c)^{-1}B_u)R^{-1/2}, \quad Q(z) = z^{-q}M(z), \quad (3.25)$$

$$A_c = (I + B_uB_u^*Z^{-1})^{-1}A_u. \quad (3.26)$$

We summarize the detailed derivation of the inner-outer factorization as follows:

Algorithm for Inner-Outer Factorization:

- Step 1: Obtain (A, B, C, D) , a realization of $T(z)$ according to (3.16) and (3.17).
- Step 2: Find a unitary matrix U such that $\bar{A} = UAU^*$ has the form in (3.20).
- Step 3: Compute the solution Z for the Lyapunov equation in (3.22).
- Step 4: Compute K_u as in (3.24) and obtain $M(z)$ and $Q(z)$ as in (3.25).
- Step 5: Compute X as in (3.23) and set $K = -(I + B^*XB)^{-1}B^*XA$.
- Step 6: Obtain the outer factor $P(z) = 1/N(z)$ where $N(z)$ has the form in (3.19).

3.3.3 COMPUTATION OF \mathcal{H}_2 NORM

Since $M(z)$ and $Q(z)$ are allpass, $M^\sim(z) = M(z^{-1}) = 1/M(z)$, and $Q^\sim(z) = Q(z^{-1}) = 1/Q(z)$. For $T_{id}(z) = cz^{-d}$ and $s = d - q > 0$, there holds

$$\begin{aligned} 2T_{id}(z)Q^\sim(z) &= 2cz^{q-d}M(z^{-1}) = 2cz^{-s} \left(D_m + C_m(z^{-1}I - A_m)^{-1}B_m \right), \\ A_m &= A_c, \quad B_m = B_u R^{-1/2}, \quad C_m = K_u, \quad D_m = R^{-1/2}, \end{aligned} \quad (3.27)$$

where A_m^{-1} is completely unstable. It follows that

$$\Pi_a [2T_{id}(z)Q^\sim(z)] = 2c \left(C_m A_m^s (z^{-1}I - A_m)^{-1} B_m \right), \quad (3.28)$$

$$\Pi_c [2T_{id}(z)Q^\sim(z)] = 2c \left(D_m z^{-s} + \sum_{i=0}^{s-1} C_m A_m^{-i+s-1} B_m z^{-i} \right), \quad (3.29)$$

The next result is obtained by a direct application of the result in [104].

Proposition 3.6 *Define Y as the solution of the Lyapunov equation*

$$A_m^* Y A_m - Y + (C_m A_m^s)^* C_m A_m^s = 0. \quad (3.30)$$

Then $\|\Pi_a [2T_{id} Q^\sim]\|_2 = 2|c| \sqrt{B_m^ Y B_m}$.*

3.3.4 COMPUTATION OF HANKEL NORM

In this subsection we consider computation of the Hankel norm of $\Pi_a [2T_{id}(z)Q^\sim(z)]$ as well as its optimal Hankel norm approximation. By (3.28), the controllability and the observability grammian of $\Pi_a [2T_{id}(z)Q^\sim(z)]$ are the solutions of two Lyapunov equations $A_m L_c A_m^* - L_c + B_m B_m^* = 0$, and $A_m^* L_o A_m - L_o + (C_m A_m^s)^* C_m A_m^s = 0$, respectively. In light of [104], there holds

$$\frac{1}{2|c|} \|\Pi_a [2T_{id} Q^\sim]\|_H = \frac{1}{2|c|} \inf_{G_o \in \mathcal{H}_\infty} \|G_o - \Pi_a [2T_{id} Q^\sim]\|_\infty = \bar{\sigma}(L_c L_o) =: \sigma_{\max} \quad (3.31)$$

where $\bar{\sigma}(\cdot)$ denotes the largest singular value. To compute the optimal approximant $G_o(z)$, let η and ξ be the Schmidt pair satisfying $L_c L_o \eta = \sigma_{\max}^2 \eta$, and $\xi = \frac{1}{\sigma_{\max}} L_o \eta$. Then the following two vectors of infinite size

$$u = \begin{bmatrix} B_m^* \\ (A_m B_m)^* \\ (A_m^2 B_m)^* \\ \vdots \end{bmatrix} \xi, \quad v = \begin{bmatrix} C_m \\ C_m A_m \\ C_m A_m^2 \\ \vdots \end{bmatrix} A_m^s \eta \quad (3.32)$$

have bounded two norm, and $U(z) = \sum_{i=1}^{\infty} B^*(A^*)^{i-1} \xi z^{-i} = B_m^* (zI - A_m^*)^{-1} \xi$, $V(z) = \sum_{i=0}^{\infty} C_m A_m^{s+i} \eta z^i = C_m A_m^s \eta + C_m A_m^{s+1} (z^{-1}I - A_m)^{-1} \eta$. Both of them are absolutely square integrable on unit circle. Note that $U(z)$ and $V(z)$ can also be

written as $U(z) = \xi^*(zI - A_m)^{-1}B_m$, $V(z) = C_m A_m^s (I + A_m(z^{-1}I - A_m)^{-1})\eta$. In this case, the optimal Hankel approximation of $G_o(z)$ is given by $G_o(z) = \Pi_a [2T_{id}Q^\sim] - 2c\sigma_{\max}V(z)/U(z)$.

An alternative for computation of optimal Hankel-norm approximant $G_o(z)$ is using the algorithm in [21] (Theorem 6.3) that was developed for continuous-time systems. Bilinear transform

$$z = \frac{\lambda - s}{\lambda + s}, \quad s = \lambda \frac{1 - z}{1 + z}, \quad \lambda > 0,$$

can be used to accomplish this goal. The following outlines the steps for using the algorithm in [21] to compute a realization (A_o, B_o, C_o, D_o) for the optimal approximant $G_o(z)$.

Alternate Algorithm for Optimal Hankel Approximation:

- Step 1: Obtain a discrete realization (A_d, B_d, C_d, D_d) for $\Pi_a [2T_{id}(z)Q^\sim(z)]$, where $A_d = A_m$, $B_d = B_m$, $C_d = 2|c|C_m A_m^s$, and $D_d = 0$.
- Step 2: Obtain a continuous realization (A_c, B_c, C_c, D_c) by using bilinear transform which gives

$$A_c = \lambda(I + A_d)^{-1}(A_d - I), \quad B_c = \sqrt{2\lambda}(I + A_d)^{-1}B_d, \quad (3.33)$$

$$C_c = \sqrt{2\lambda}C_d(I + A_d)^{-1}, \quad D_c = D_d - C_d(I + A_d)^{-1}B_d. \quad (3.34)$$

- Step 3: Compute the balanced realization (A, B, C, D) using the algorithm in [21] which yields $AP + PA^* + BB^* = 0$, $A^*Q + QA + C^*C = 0$, $P = Q = \text{diag}(\sigma I_r, \Sigma)$, and $\sigma I - \Sigma > 0$.

- Step 4: Partition (A, B, C) conformally with P , as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & C_2 \end{bmatrix},$$

and define

$$\hat{A} = \Gamma^{-1}(\sigma^2 A_{22}^* + \Sigma A_{22} \Sigma - \sigma C_2^* U B_2^*), \quad (3.35)$$

$$\hat{B} = \Gamma^{-1}(\Sigma B_2 + \sigma C_2^* U), \quad (3.36)$$

$$\hat{C} = C_2 \Sigma + \sigma U B_2^*, \quad (3.37)$$

$$\hat{D} = D - \sigma U, \quad (3.38)$$

where U is a unitary matrix satisfying $B_1 = -C_1^* U$ and $\Gamma = \Sigma^2 - \sigma^2 I$.

- Step 5: Transform the continuous realization $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ back to a discrete realization using the bilinear transform:

$$A_o = (\lambda I - \hat{A})(\lambda I + \hat{A})^{-1}, \quad B_o = \sqrt{2\lambda}(\lambda I + \hat{A})^{-1} \hat{B}, \quad (3.39)$$

$$C_o = -\sqrt{2\lambda} \hat{C}(\lambda I + \hat{A})^{-1}, \quad D_o = \hat{D} - \hat{C}(\lambda I + \hat{A})^{-1} \hat{B}. \quad (3.40)$$

3.4 DESIGN EXAMPLES

Two design examples will be used to illustrate the design algorithms proposed in this chapter. To simplify the numerical computation, we assume that $H_1(z) = H_0(-z)$, although this is not required in using our proposed design method. Actually just as what [90] pointed out, $H_1(z) = H_0(-z)$ should be avoided in order to achieve perfect reconstruction. However, we want to consider the case when perfect

reconstruction is not a realistic goal to achieve. By assuming minimal realization for $F_0(z)$ and $F_1(z)$, for FIR analysis filter banks, there hold

$$\text{order } F_0 = d + \text{order } H_1 - q, \quad \text{order } F_1 = d + \text{order } H_0 - q,$$

for the \mathcal{H}_2 optimal approximation algorithm where normally $q = 1$, and

$$\text{order } F_0 = d + \text{order } H_1 - q + \kappa - 1, \quad \text{order } F_1 = d + \text{order } H_0 - q + \kappa - 1,$$

for the \mathcal{H}_∞ optimal approximation algorithm where κ is the number of unstable zeros of $\det H_{AC}$. For IIR analysis filter banks, there hold

$$\text{order } F_0 = \text{order } F_1 = d + 2(\text{order } H_0 + \text{order } H_1 - q),$$

for the \mathcal{H}_2 optimal approximation algorithm, and

$$\text{order } F_0 = d + \text{order } H_1 - q + \kappa - 1, \quad \text{order } F_1 = d + \text{order } H_0 - q + \kappa - 1,$$

for the \mathcal{H}_∞ optimal approximation algorithm. It can be seen that for both FIR and IIR analysis filter banks, the order of F_0 and F_1 satisfy the same corresponding formula by using the \mathcal{H}_∞ optimal approximation algorithm. Because large d is often required in order to achieve small reconstruction error, model reduction is necessary. For the two examples in this section, balanced truncation is employed for model reduction of $G(z)$ (in stead of $F_0(z)$ and $F_1(z)$ directly) to ensure zero aliasing error. It should also be pointed out that the orders of F_0 and F_1 obtained using our proposed algorithms are smaller than those using the algorithm in [3] because of the

use of polyphase decomposition in [3] that tends to double the realization of $G_o(z)$. Also we need to point out that the percentage reconstruction error in our examples is defined as

$$e\% = \left\| \frac{1}{2c} G(z) \det H_{Ac}(z) - z^{-d} \right\|_{\infty} \times 100$$

Example 3.1: FIR analysis filter bank.

In this example, $H_0(z)$ is a 19th order half band low pass filter designed using the Remez algorithm with transition band $[.45\pi, .55\pi]$ which can be easily computed with MATLAB command *remez*. $H_1(z)$ is taken to be $H_0(-z)$. These two filters are not power complementary. So the perfect reconstruction does not apply [90]. The magnitude Bode plots of these two filters are shown in Figure 3.2. For the specification of 2% reconstruction error over all frequency range, \mathcal{H}_{∞} optimization is required that yields time delay $d = 39$. The corresponding optimal synthesis filters are computed that have a minimal order 75. Because the order is quite high, model reduction is carried out for $G(z)$ that reduces the order of $F_0(z)$ and $F_1(z)$ to 39 without notable distortion for a 1.8% reconstruction error. The Bode magnitude plots of reduced synthesis filters \hat{F}_0 and \hat{F}_1 are shown in Figure 3.3. The magnitude frequency response of the associated reconstruction error and zero aliasing error are shown in Figure 3.4 with solid lines.

Example 3.2: IIR analysis filter bank.

This example is the same as Example 3 of [3] where $H_0(z)$ is the bilinear transformation of a fifth-order analog elliptic filter (without prewarping) with cutoff frequency 1 rad/s. Thus $H_0(z)$ is a fifth-order lowpass filter with cutoff frequency $\pi/2$. Taking $H_1(z) = H_0(-z)$ implies that $H_1(z)$ is a fifth-order highpass filter. Figure 3.5 shows the magnitude Bode plots of $H_0(z)$ and $H_1(z)$. With the same 2%

specification for the reconstruction error over all frequency range, the time delay $d = 47$ is required when using \mathcal{H}_∞ optimization algorithm proposed in this paper. The orders of $F_0(z)$ and $F_1(z)$ are 68, compared to the order of 101 in [3]. Again, model reduction is carried out to $G(z)$ that reduces the orders of the synthesis filters to 55 without notable distortion for the reconstruction error. The magnitude Bode plots of the reduced synthesis filters $\hat{F}_0(z)$ and $\hat{F}_1(z)$ are shown in Figure 3.6. The associated reconstruction error is less than 2% as shown in Figure 3.7 with solid line. With balanced truncation for $G(z)$, the lowest order of $\hat{F}_0(z)$ and $\hat{F}_1(z)$ that can be reduced to is 51 without violating the 2% specification for the reconstruction error. The associated reconstruction error is shown in Figure 3.7 with a dashed line. It is seen that the only distortion in this case occurs in the transition band range.

The model matching method yields good design for the two-channel multirate filter banks. It can also be extended to multi-channel filter bank design which is the topic of the next chapter.

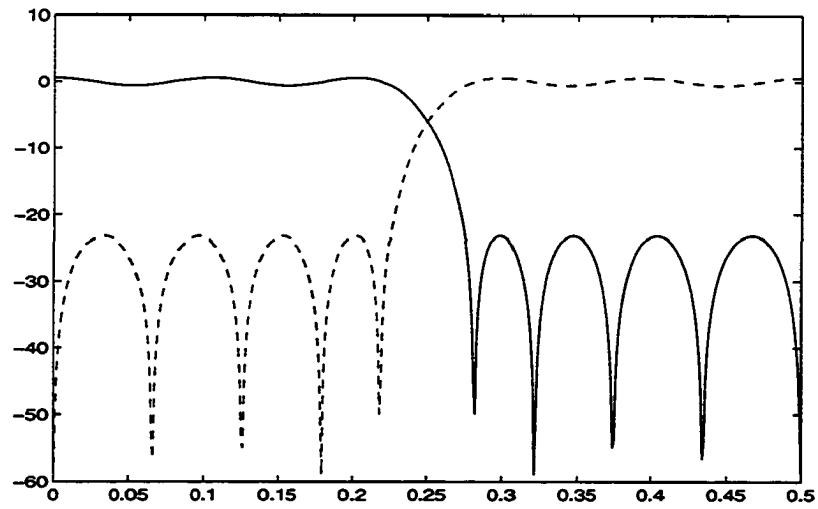


Figure 3.2: Example 3.1: $|H_0|$ and $|H_1|$ in dB vs. $\omega/2\pi$.

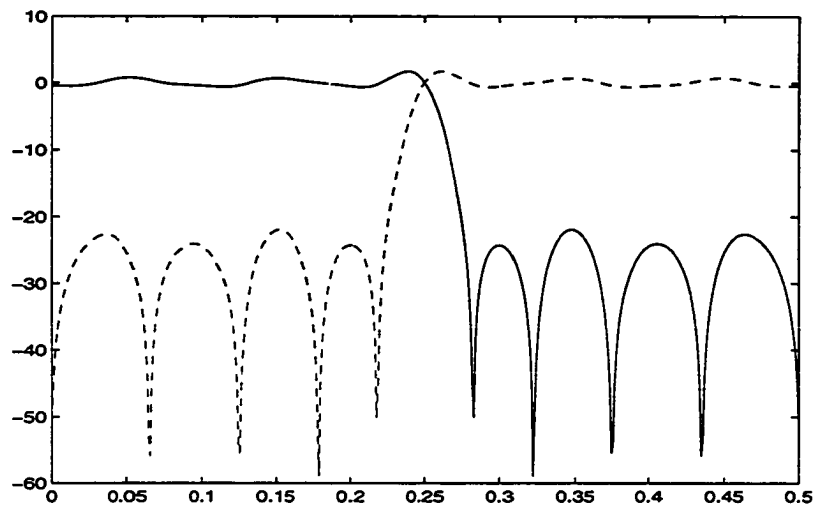


Figure 3.3: Example 3.1: $|\hat{F}_0|$ and $|\hat{F}_1|$ in dB vs. $\omega/2\pi$.

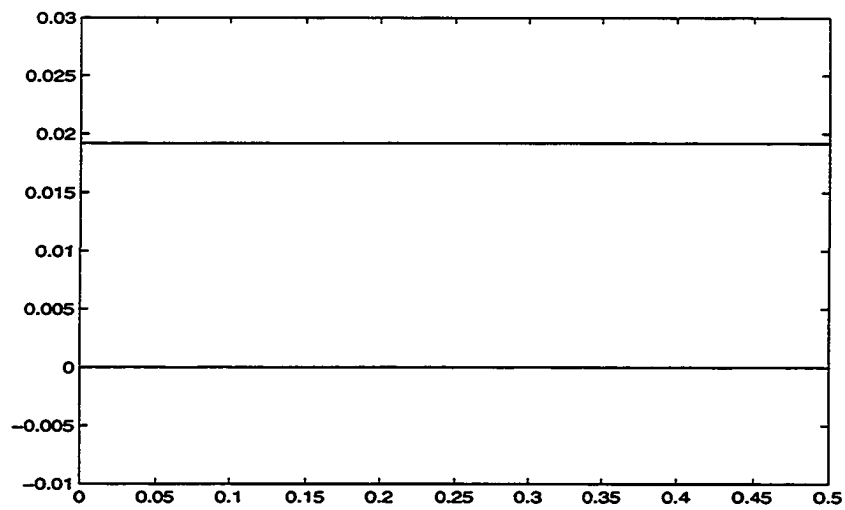


Figure 3.4: Example 3.1: Reconstruction and aliasing error vs. $\omega/2\pi$.

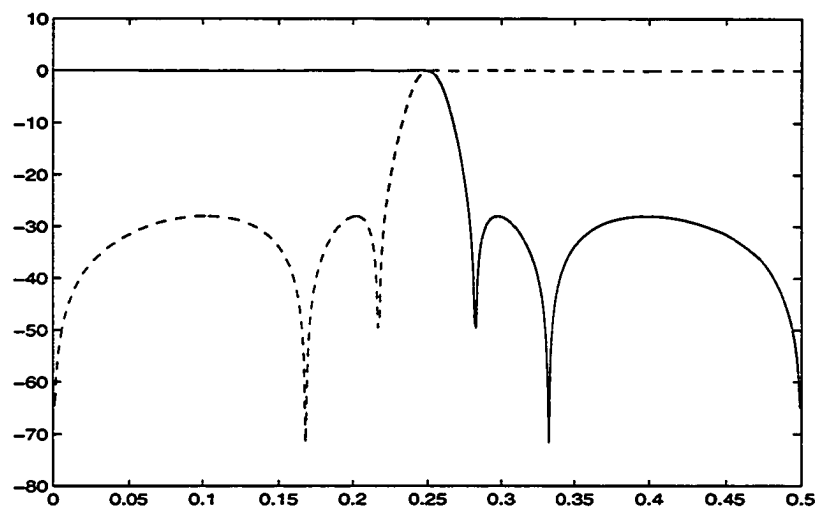


Figure 3.5: Example 3.2: $|H_0|$ and $|H_1|$ in dB vs. $\omega/2\pi$.

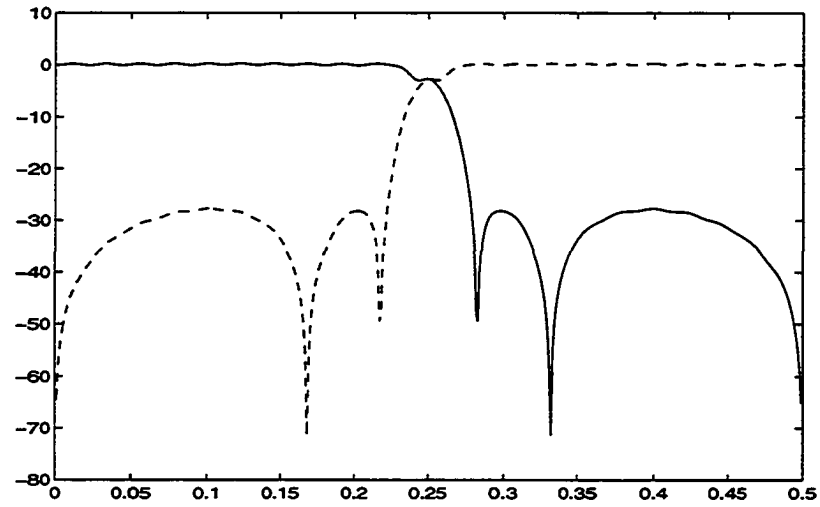


Figure 3.6: Example 3.2: $|\hat{F}_0|$ and $|\hat{F}_1|$ in dB vs. $\omega/2\pi$.

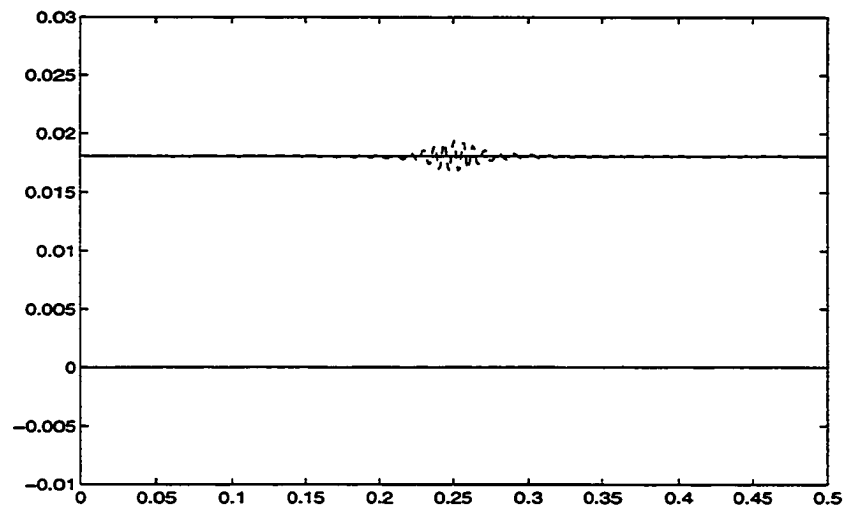


Figure 3.7: Example 3.2: Reconstruction and aliasing error vs. $\omega/2\pi$.

CHAPTER 4

MULTI-CHANNEL FILTER BANK DESIGN AND ANALYSIS

4.1 MULTIRATE SYSTEMS TO BE DESIGNED

This chapter will generalize the results in the previous chapter to multi-channel filter banks. We first consider the M -channel maximally decimated filter bank [90] as shown in Figure 4.1. This kind of uniform filter bank is also known as the critical sampled filter bank because the decimation ratio M is the same as the number of channels. The analysis and synthesis filter bank consist of, respectively, $H_i(z)$ and $F_i(z)$, $i = 1, 2, \dots, M - 1$, which are restricted to be causal and stable transfer functions of linear time-invariant filters. The input signal $x(n)$ is split into M frequency subbands by M analysis filters $H_i(z)$ according to the energy distribution of the signal $x(n)$ in frequency domain. Each subband signal $x_i(n)$, $i = 1, 2, \dots, M - 1$, is down-sampled by a factor of M , then coded and transmitted. We assume that the quantization error of coding and transmission is negligible. In the receiver end, each subband signal is up-sampled by a factor of M . After passing through M synthesis filters $F_i(z)$ respectively, the M subband signals, $\hat{x}_i(n)$, $i = 1, 2, \dots, M - 1$, are combined to form the reconstructed signal $\hat{x}(n)$. This kind of system has important application value in subband coding of speech and digital audio; See, for example, [58, 60, 61, 63, 65, 88] and their corresponding references.

In the following sections, by using the model matching method, an error system is set up as shown in Figure 4.2. The design issue will be investigated for multi-channel filter banks that achieve zero aliasing error, and minimize certain measure

of the error signal $e(n) = \hat{x}(n) - x(n - d)$ for some $d \geq 0$. This will lead to the extension of \mathcal{H}_2 and \mathcal{H}_∞ optimization techniques to the design of synthesis filter banks. Optimal solutions will be derived and numerical algorithms for computing the optimal synthesis filter banks will be developed. The advantage of our proposed approach is the simplicity and efficiency of the design algorithm, as well as zero aliasing error that is contrast to [74, 3]. It is known that nonuniform multirate filter banks, i.e, all channels in Figure 4.1 do not have the same up and down sampling rate, can be transformed into uniform ones [43] while uniform filter banks can also be used to generate nonuniform ones [59]. Therefore, the results of this chapter regarding unifrom filter banks have applications to the design of nonuniform filter banks.

4.2 DESIGN OF SYNTHESIS FILTER BANKS WITH \mathcal{H}_2 AND \mathcal{H}_∞ OPTIMIZATIONS

4.2.1 FORMULATION OF THE DESIGN APPROACH

Consider the maximally decimated filter bank in Figure 4.1 where all the filters are restricted to be both causal and stable. This multirate filter bank is generalized from the two-channel filter bank introduced in the previous two chapters. Compared with the case of two-channel filter bank, the input/output relation of this M -channel filter bank is more complicated. In \mathcal{Z} -transform, it is given by

$$\hat{X}(z) = \sum_{l=0}^{M-1} A_l(z)X(zW^l), \quad A_l(z) = \frac{1}{M} \sum_{k=0}^{M-1} H_k(zW^l)F_k(z) \quad (4.1)$$

where $0 \leq l \leq M-1$ and $W = e^{-j2\pi/M}$. Those terms involving $X(zW^l)$ for $l \geq 1$ are called aliasing error. If the synthesis filters $F_i(z)$, $i = 1, 2, \dots, M-1$, are properly

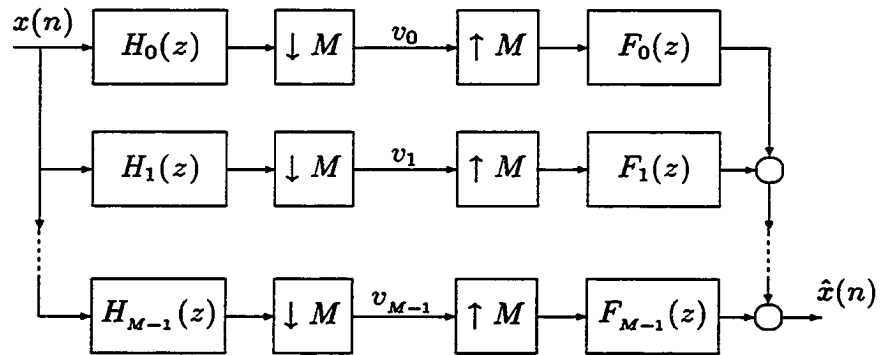


Figure 4.1: M -channel maximally decimated filter bank.

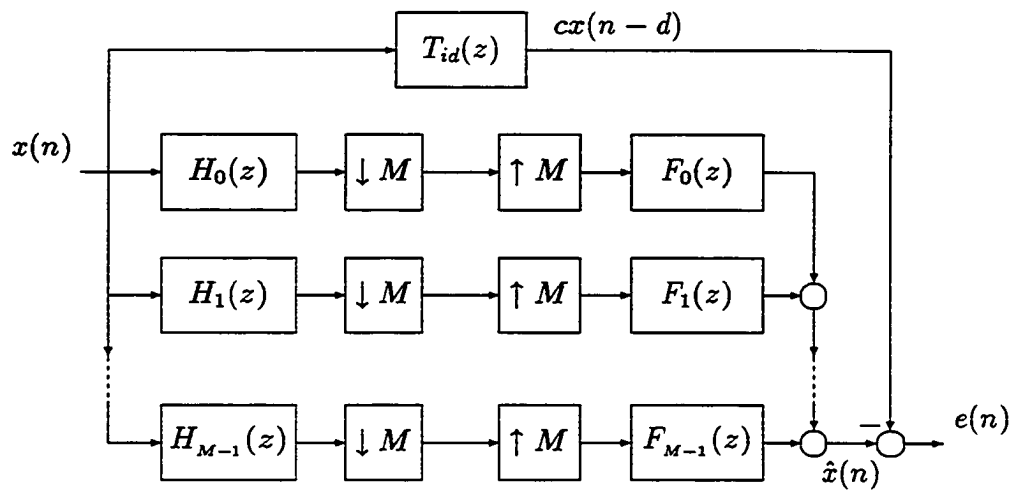


Figure 4.2: Multi-channel error system.

chosen, then the aliasing error will be eliminated completely, that is, $A_l(z) = 0$ for $1 \leq l \leq M-1$ and

$$\hat{X}(z) = T_r(z)X(z) = A_0(z)X(z), \quad T_r(z) = A_0(z) = \frac{1}{M} \sum_{k=0}^{M-1} H_k(z)F_k(z). \quad (4.2)$$

But if $T_r(z) \neq cz^{-d}$ where $c \neq 0$ and $d \geq 0$, then $\hat{X}(z)$ may involve either amplitude distortion or phase distortion or both. We introduce the following definition regarding perfect reconstruction (PR) given by [90].

Definition 4.1 *A maximally decimated filter bank shown in Figure 4.1 is said to have perfect reconstruction (PR) property if it is free from aliasing, amplitude distortion and phase distortion, i.e.,*

$$\hat{X}(z) = T_r(z)X(z), \quad T_r(z) = \frac{1}{M} \sum_{k=0}^{M-1} H_k(z)F_k(z) = cz^{-d}, \quad (4.3)$$

for some $c \neq 0$ and $d \geq 0$.

The next theorem follows easily from the above definition, and therefore the proof is skipped.

Theorem 4.1 *The maximally decimated filter bank shown in Figure 4.1 has the perfect reconstruction (PR) property if and only if*

$$\frac{1}{M} \begin{bmatrix} H_0(z) & H_1(z) & \cdots & H_{M-1}(z) \\ H_0(zW) & H_1(zW) & \cdots & H_{M-1}(zW) \\ \vdots & \vdots & \vdots & \vdots \\ H_0(zW^{M-1}) & H_1(zW^{M-1}) & \cdots & H_{M-1}(zW^{M-1}) \end{bmatrix} \begin{bmatrix} F_0(z) \\ F_1(z) \\ \vdots \\ F_{M-1}(z) \end{bmatrix} = \begin{bmatrix} cz^{-d} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

for some $c \neq 0$ and $d \geq 0$.

Define the $M \times M$ aliasing component matrix $H_{AC}(z)$ and the $M \times 1$ synthesis matrix $F(z)$ as

$$H_{AC}(z) = \begin{bmatrix} H_0(z) & H_1(z) & \cdots & H_{M-1}(z) \\ H_0(zW) & H_1(zW) & \cdots & H_{M-1}(zW) \\ \vdots & \vdots & \vdots & \vdots \\ H_0(zW^{M-1}) & H_1(zW^{M-1}) & \cdots & H_{M-1}(zW^{M-1}) \end{bmatrix} F(z) = \begin{bmatrix} F_0(z) \\ F_1(z) \\ \vdots \\ F_{M-1}(z) \end{bmatrix} \quad (4.4)$$

Define the $1 \times M$ analysis matrix $H(z)$ as

$$H(z) = \begin{bmatrix} H_0(z) & H_1(z) & \cdots & H_{M-1}(z) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} H_{AC}(z) \quad (4.5)$$

which is the first row of $H_{AC}(z)$. If $\det H_{AC}(z) \neq 0$ for all $|z| \geq 1$, then with $c = \frac{1}{M}$ and $d = 0$, the ideal synthesis matrix

$$F_{id}(z) := H_{AC}^{-1}(z) \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \frac{\text{adj}(H_{AC}(z))}{\det H_{AC}(z)} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (4.6)$$

is causal, stable, and achieves perfect reconstruction. However, if $\det H_{AC}(z)$ has unstable zeros, then the ideal synthesis matrix $F_{id}(z)$ as in (4.6) is either noncausal or unstable or both. Because noncausal and/or unstable filters can not be used for real time implementation, perfect reconstruction is not possible for the multirate filter bank. In this case we want to seek a synthesis matrix $F(z)$ which will minimize either the covariance or the energy of the error signal $e(n)$ shown in Figure 4.2

subject to the constraint of zero aliasing error. We will show that it actually leads to equivalent frequency domain optimization problems of \mathcal{H}_2 and \mathcal{H}_∞ . The design algorithms for the optimal synthesis filter banks will be developed subsequently. According to [90], the polyphase decomposition of the transfer function of the k th analysis filter $H_k(z)$ is given by $H_k(z) = \sum_{l=0}^{M-1} z^{-l} E_{k,l}(z^M)$. So the analysis matrix has the expression

$$\begin{aligned} H(z) &= \begin{bmatrix} H_0(z) & H_1(z) & \cdots & H_{M-1}(z) \end{bmatrix} \\ &= \begin{bmatrix} 1 & z^{-1} & \cdots & z^{-(M-1)} \end{bmatrix} E^T(z^M), \end{aligned} \quad (4.7)$$

$$E(z) = \begin{bmatrix} E_{0,0}(z) & E_{0,1}(z) & \cdots & E_{0,M-1}(z) \\ E_{1,0}(z) & E_{1,1}(z) & \cdots & E_{1,M-1}(z) \\ \vdots & \vdots & \cdots & \vdots \\ E_{M-1,0}(z) & E_{M-1,1}(z) & \cdots & E_{M-1,M-1}(z) \end{bmatrix}, \quad (4.8)$$

$$\begin{aligned} H(zW^i) &= \begin{bmatrix} 1 & W^{-i}z^{-1} & \cdots & W^{-(M-1)i}z^{-(M-1)} \end{bmatrix} E^T(z^M) \\ &= \begin{bmatrix} 1 & W^{-i} & \cdots & W^{-(M-1)i} \end{bmatrix} \text{diag} \left\{ z^{-k} \right\}_{k=0}^{M-1} E^T(z^M) \\ &= \begin{bmatrix} 1 & W^{-i} & \cdots & W^{-(M-1)i} \end{bmatrix} \tilde{H}_{AC}(z), \end{aligned} \quad (4.9)$$

for $i = 0, 1, \dots, M-1$ with $\tilde{H}_{AC}(z) = \text{diag} \left\{ z^{-k} \right\}_{k=0}^{M-1} E^T(z^M)$. Now the alias component matrix as in (4.4) can be written into a product form $H_{AC}(z) = L\tilde{H}_{AC}(z)$, where L is a (complex) constant matrix given by

$$L = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & W^{-1} & \cdots & W^{-(M-1)} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & W^{-(M-1)} & \cdots & W^{-(M-1)(M-1)} \end{bmatrix}$$

The next result shows that the desired synthesis matrix $F(z)$ shown in Figure 4.2 can be derived from one compact form with the fact that the determinant of $H_{AC}(z)$ satisfies

$$\det H_{AC}(z) = (\det L)(\det \tilde{H}_{AC}(z)) = (\det L)z^{-(M-1)M/2} \det E(z^M). \quad (4.10)$$

Theorem 4.2 *Suppose that the analysis filter bank $H(z)$ is causal and stable for the maximally decimated filter bank shown in Figure 4.1. Then all causal and stable synthesis filter banks $F(z)$ achieving zero aliasing error are parameterized as*

$$F(z) = G(z)(\det L)^{-1} \text{adj}(H_{AC}(z)) \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}^T \quad (4.11)$$

for some causal and stable $G(z)$ that is real valued for each real z .

Proof: Any causal and stable synthesis filter bank $F(z)$ achieving zero aliasing error satisfies $\begin{bmatrix} 0 & I_{M-1} \end{bmatrix} H_{AC}(z) F(z) = 0$, where I_k is an identity matrix of size k . Since $H(z)$ is causal and stable, $H(z)F(z) = \phi(z)$ is causal and stable as well. Combining these two facts yields

$$F(z) = H_{AC}^{-1}(z) \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}^T \phi(z) = F_{id}(z)\phi(z),$$

with $F_{id}(z)$ as in (4.6). By the fact that $H(z)F_{id}(z) = 1$, and $H(z)$ is stable and causal, $F_{id}(z)$ has no unstable block zero. It follows from the causality and stability of $F(z)$ that

$$\phi(z) = G(z)(\det L)^{-1} \det H_{AC}(z)$$

for some $G(z)$ which is also stable and causal. That is, the unstable poles of $F_{id}(z)$ that are also unstable zeros of $\det H_{AC}(z)$ have to be cancelled by unstable zeros of

$\phi(z)$, in order to ensure that $F(z)$ be causal and stable. Substituting the expression of $\phi(z)$ into the expression of $F(z)$ gives

$$\begin{aligned} F(z) &= F_{id}(z)G(z)(\det L)^{-1} \det H_{AC}(z) \\ &= G(z)(\det L)^{-1} \text{adj}(H_{AC}(z)) \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}^T \end{aligned}$$

which is the same as (4.11). The fact that $H(z)F_{id}(z) = 1$ and identity (4.10) imply that

$$H(z)F(z) = G(z)(\det L)^{-1} \det H_{AC}(z) = G(z)z^{-(M-1)M/2} \det E(z^M)$$

which is real valued for each real z , if, and only if $G(z)$ is. ■

Parameterization of all causal and stable synthesis filter banks achieving zero aliasing error is important. It leads to elegant optimization problems for the minimization of reconstruction error to be shown next in detail. Define $E(z) = \mathcal{Z}(e(n))$, the \mathcal{Z} -transform of the error signal $e(n)$ as in Figure 4.2. Then $E(z) = T_E(z)X(z)$, where

$$\begin{aligned} T_E(z) &= T_r(z) - T_{id}(z) = \frac{1}{M}H(z)F(z) - cz^{-d} \\ &= \frac{1}{M}G(z)z^{-(M-1)M/2} \det E(z^M) - cz^{-d}. \end{aligned} \quad (4.12)$$

Thus the aliasing error is completely eliminated. The fact that $T_E(z)$ is a scalar function yields much simpler optimization problem and more efficient design algorithms than those studied in [74, 3]. Note that for $M = 2$, (4.12) reduces to what was studied in the previous chapter.

Motivated by the model matching method in [74, 3], we seek a causal and stable transfer function $G(z)$ such that $\|T_E\|_2$ and $\|T_E\|_\infty$ are minimized. The synthesis filter bank can then be obtained according to (4.11). We will derive solutions of $G(z)$, and thus of synthesis matrix $F(z)$ in (4.11), for both \mathcal{H}_2 and \mathcal{H}_∞ optimizations.

Theorem 4.3 *Suppose that the analysis filter bank $H(z)$ is causal and stable for the maximally decimated filter bank shown in Figure 4.1. Then for each given $d \geq 0$ and $c \neq 0$, the optimal synthesis filter bank $F(z)$ that minimizes the reconstruction error $\|T_E\|_\alpha$, and achieves zero aliasing error is unique for $\alpha = 2$ and $\alpha = \infty$, where T_E is given in (4.12).*

Proof: Since all causal and stable synthesis filter banks are parameterized by (4.11), and the transfer function for the reconstruction error given in (4.12) is scalar, the optimal solution of $G(z)$ that minimizes $\|T_E\|_\alpha$ is unique for $\alpha = 2$ and $\alpha = \infty$ [104] which concludes the uniqueness of the optimal synthesis filter banks. ■

It should be clear that the integer value of d determines the reconstruction error, and the optimal reconstruction error measured in \mathcal{H}_2 or \mathcal{H}_∞ norm is a function of d . It will be proven later that as $d \rightarrow \infty$, the reconstruction error tends to zero as well. Hence PR can be achieved asymptotically with respect to time delay d . Let $(M-1)M/2 = \beta M + \gamma_M$, $d = \tilde{d}M + \gamma_d$, and $\gamma_d = \gamma_M = \gamma$, where β , γ_M , γ_d , and \tilde{d} are all integers, and $\gamma_M, \gamma_d \in \{0, 1, \dots, M-1\}$. By imposing $\gamma_d = \gamma_M = \gamma$, $G(z)$ can be taken as $G(z) = \tilde{G}(z^M)$, and thus substitute it into (4.12) yields

$$\begin{aligned} T_E(z) &= \left(\frac{1}{M} \tilde{G}(z^M) z^{-\beta M} \det E(z^M) - cz^{-\tilde{d}M} \right) z^{-\gamma} = \tilde{T}_E(z^M) z^{-\gamma} \\ \tilde{T}_E(z) &= \frac{1}{M} \tilde{G}(z) z^{-\beta} \det E(z) - cz^{-\tilde{d}}. \end{aligned} \quad (4.13)$$

Since the order of \tilde{T}_E is smaller than that of T_E by a factor of $1/M$, minimizing $\|T_E\|_\alpha$ can be done through minimizing $\|\tilde{T}_E\|_\alpha$ for $\alpha = 2, \infty$. This greatly reduces computation effort in the design process, and enhances the accuracy for large M .

4.2.2 OPTIMAL \mathcal{H}_2 AND \mathcal{H}_∞ SOLUTIONS FOR THE SYNTHESIS FILTER BANK

Consider $\tilde{T}_E(z)$ given in (4.13) where $z^{-\beta} \det E(z)$ has exactly κ zeros outside of the unit circle, counting multiplicity. Then there exists a factorization

$$z^{-\beta} \det E(z) = P(z)Q(z), \quad Q(z) = z^{-q} \frac{q_0 + q_1 z + \cdots + q_\kappa z^\kappa}{q_0 z^\kappa + q_1 z^{\kappa-1} + \cdots + q_\kappa z}, \quad q \geq 0, \quad (4.14)$$

which is the same as that in the previous chapter. Because $Q(z)Q(z^{-1}) \equiv 1$,

$$\|T_E\|_\alpha = \frac{1}{M} \|\tilde{G}PQ - M\tilde{T}_{id}\|_\alpha = \frac{1}{M} \|\tilde{G}P - M\tilde{T}_{id}Q^\sim\|_\alpha, \quad \alpha = 2, \infty, \quad (4.15)$$

where $Q^\sim = Q(z^{-1})$, and $\tilde{T}_{id}(z) = cz^{-\bar{d}}$ which is the ideal transfer function.

Theorem 4.4 *Consider the error function $\|T_E\|_\alpha$ as in (4.15) where $\alpha = 2, \infty$. $\tilde{G}(z)$ is causal and stable to be determined. Let $z^{-\beta} \det E(z) = P(z)Q(z)$ be the inner-outer factorization as in (4.14) where $Q(z)$ includes all the unstable zeros of $z^{-\beta} \det E(z)$. Denote \tilde{G}_2 and \tilde{G}_∞ as the optimal \mathcal{H}_2 and \mathcal{H}_∞ solutions of $\tilde{G}(z)$. Then with $\tilde{T}_{id}(z) = cz^{-\bar{d}}$ the \mathcal{H}_2 and \mathcal{H}_∞ optimal solutions are respectively given by*

$$\tilde{G}_2(z) = \Pi_c [M\tilde{T}_{id}(z)Q^\sim(z)] / (P(z)), \quad (4.16)$$

$$\tilde{G}_\infty(z) = \Pi_c [M\tilde{T}_{id}(z)Q^\sim(z)] / P(z) \quad (4.17)$$

$$+ \operatorname{argmin} \left\{ \tilde{G}_o(z) : \inf_{\tilde{G}_o \in \mathcal{H}_\infty} \|\tilde{G}_o - \Pi_a [M\tilde{T}_{id}Q^\sim]\|_\infty \right\} / P(z),$$

$$\inf_{\tilde{G} \in \mathcal{H}_2} \|T_E\|_2 = \|\Pi_a [\tilde{T}_{id}Q^\sim]\|_2, \quad \inf_{\tilde{G} \in \mathcal{H}_\infty} \|T_E\|_\infty = \|\Pi_a [\tilde{T}_{id}Q^\sim]\|_\infty.$$

Theorem 4.5 *Suppose that $\det E(z)$ avoids zeros on the unit circle. Then with the optimal \mathcal{H}_2 and \mathcal{H}_∞ solutions in Theorem 4.4, the reconstruction error approaches to zero as $\tilde{d} \rightarrow \infty$.*

The above two results are parallel to Theorem 3.3 and Theorem 3.4 respectively so that the proof is reasonably skipped.

4.2.3 DESIGN ALGORITHMS FOR THE SYNTHESIS FILTER BANK

The results in the previous two subsections yield the following two iterative algorithms for the design of synthesis filter banks based on \mathcal{H}_2 and \mathcal{H}_∞ optimizations. We again assume that $\tilde{T}_{id}(z) = cz^{-\tilde{d}}$ where $c \neq 0$ and $\tilde{d} > 0$ which is to be determined in the design process. The first algorithm is based on \mathcal{H}_2 optimization and the second based on \mathcal{H}_∞ optimization. Note that for $M = 2$, rather than compute $z^{-\beta} \det E(z)$, $\det H_{AC}(z)$ will be computed directly.

Design Algorithm 1 (\mathcal{H}_2 Optimization):

- Step 1: For the given analysis filters $H_0(z), H_1(z), \dots, H_{M-1}(z)$ and the design specification ϵ_2 , compute $z^{-\beta} \det E(z)$.
- Step 2: Compute inner-outer factorization $z^{-\beta} \det E(z) = P(z)Q(z)$ where $Q(z)$ is an allpass function of the form in (4.14) that contains all unstable zeros of $\det E(z)$.
- Step 3: Compute the optimal \mathcal{H}_2 performance index: $M \|\Pi_a [\tilde{T}_{id} Q^\sim]\|_2$.
- Step 4: If $M \|\Pi_a [\tilde{T}_{id} Q^\sim]\|_2 > \epsilon_2$, then increase \tilde{d} and go to Step 3; Otherwise go to Step 5.
- Step 5: Compute $\tilde{G}(z) = G_2(z) = \Pi_c [M \tilde{T}_{id}(z) Q^\sim(z)] / P(z)$ and set synthesis filter bank according to (4.11).

Design Algorithm 2 (\mathcal{H}_∞ Optimization):

- Step 1: For the given analysis filters $H_0(z), H_1(z), \dots, H_{M-1}(z)$ and the design specification ϵ_∞ , compute $z^{-\beta} \det E(z)$.
- Step 2: Compute inner-outer factorization $z^{-\beta} \det E(z) = P(z)Q(z)$ where $Q(z)$ is allpass function of the form in (4.14) that contains all unstable zeros of $\det E(z)$.
- Step 3: Compute the optimal \mathcal{H}_∞ performance index: $M \|\Pi_a [\tilde{T}_{id} Q^\sim]\|_H$.
- Step 4: If $M \|\Pi_a [\tilde{T}_{id} Q^\sim]\|_H > \epsilon_\infty$, then increase d and go to Step 3; Otherwise go to Step 5.
- Step 5: Compute $\tilde{G}(z) = \tilde{G}_\infty(z)$ according to (4.17) and set synthesis filter bank according to (4.11).

Two comments are in order. First in the above two algorithms, the search of the smallest \tilde{d} value for the optimal reconstruction error to satisfy the given specification is rather primitive. If bisection or other more efficient method is employed [80], it will speed up the computation of the right \tilde{d} value. Second, although the two algorithms seem straightforward, their numerical implementation with good numerical property and efficiency is not that simple. In the next section we will study the state-space computational issue which makes the two algorithms work more efficiently.

4.3 STATE-SPACE COMPUTATION

Despite the similarity to the previous chapter for the \mathcal{H}_2 and \mathcal{H}_∞ optimization, there are differences in the numerical implementation. Thus we will elaborate in more detail. We begin with state-space realization for matrix $E^T(z^M)$. By assuming

that $H_k(z)$ shares the same denominator for $k = 0, 1, \dots, M-1$, a simple realization can be found such that

$$H_k(z) = \sum_{i=0}^{\infty} H_{k,i} z^{-i} = D_k + \sum_{i=1}^{\infty} C A^{i-1} B_k z^{-i}.$$

This is true for FIR analysis filter banks, and can also be made true for IIR case.

The sequence $\{H_{k,i}\}_{i=0}^{\infty}$ is the impulse response of $H_k(z)$ with

$$H_{k,i} = \begin{cases} D_k, & i = 0, \\ C A^{i-1} B_k, & i > 0. \end{cases} \quad (4.18)$$

Since the index i can be written as $i = sM + l$ for $l = 0, 1, \dots, M-1$, and $s = 0 \rightarrow \infty$,

$$H_k(z) = \sum_{l=0}^{M-1} \sum_{s=0}^{\infty} H_{k,sM+l} z^{-(sM+l)} = \sum_{l=0}^{M-1} z^{-l} E_{k,l}(z), \quad E_{k,l}(z^M) = \sum_{s=0}^{\infty} H_{k,sM+l} (z^M)^{-s}.$$

Clearly $E_{k,l}(z^M)$ is the l th polyphase component of $H_k(z)$. For $l = 0$,

$$\begin{aligned} E_{k,0}(z^M) &= D_k + \sum_{s=1}^{\infty} C A^{sM-M+M-1} B_k (z^M)^{-s} \\ &= D_k + C A^{M-1} (z^M I - A^M)^{-1} B_k, \end{aligned}$$

by (4.18). Similarly for $l > 0$,

$$\begin{aligned} E_{k,l}(z^M) &= \sum_{s=0}^{\infty} C A^{sM-M+M+l-1} B_l (z^M)^{-s} \\ &= C A^{l-1} B_k + \sum_{s=1}^{\infty} C A^{sM-M+M+l-1} B_k (z^M)^{-s} \\ &= C A^{l-1} B_k + C A^{M+l-1} \sum_{s=1}^{\infty} (A^M)^{s-1} B_k (z^M)^{-s} \\ &= C A^{l-1} B_k + C A^{M+l-1} (z^M I - A^M)^{-1} B_k. \end{aligned}$$

Therefore, the polyphase components with respect to the analysis matrix has an expression

$$E_{k,l}(z) = \begin{cases} D_k + CA^{M-1}(zI - A^M)^{-1}B_k, & l = 0, \\ CA^{l-1}B_k + CA^{M+l-1}(zI - A^M)^{-1}B_k, & l > 0. \end{cases}$$

It follows that

$$H(z) = \begin{bmatrix} 1 & z^{-1} & \dots & z^{-(M-1)} \end{bmatrix} \begin{bmatrix} E_{0,0}(z^M) & E_{1,0}(z^M) & \dots & E_{M-1,0}(z^M) \\ E_{0,1}(z^M) & E_{1,1}(z^M) & \dots & E_{M-1,1}(z^M) \\ \vdots & \dots & \dots & \vdots \\ E_{0,M-1}(z^M) & E_{1,M-1}(z^M) & \dots & E_{M-1,M-1}(z^M) \end{bmatrix},$$

$$\begin{bmatrix} E_{k,0} \\ E_{k,1} \\ \vdots \\ \vdots \\ E_{k,M-1} \end{bmatrix} = \begin{bmatrix} D_k \\ CB_k \\ CAB_k \\ \vdots \\ CA^{M-2}B_k \end{bmatrix} + \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{M-1} \end{bmatrix} A^{M-1}(z^M I - A^M)^{-1}B_k.$$

A state-space realization for $E^T(z)$ can now be obtained as

$$E^T(z) = \begin{bmatrix} D \\ CB \\ CAB \\ \vdots \\ CA^{M-2}B \end{bmatrix} + \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{M-1} \end{bmatrix} A^{M-1}(zI - A^M)^{-1}B,$$

where

$$B = \begin{bmatrix} B_0 & B_1 & \dots & B_{M-1} \end{bmatrix}, \quad D = \begin{bmatrix} D_0 & D_1 & \dots & D_{M-1} \end{bmatrix}.$$

With the realization for $E^T(z)$, a realization for $z^{-\beta} \det E(z)$ can also be found which can be written as

$$z^{-\beta} \det E(z) = C_E(zI - A_E)^{-1} B_E. \quad (4.19)$$

If minimal realization is preferred, then balance truncation [104] can be employed to eliminate the uncontrollable and unobservable modes.

The numerical computation of inner-outer factorization for (4.19) can be implemented by using the same algorithm as in Chapter 3. The same is also true for numerical computation of \mathcal{H}_2 and Hankel norm associated with the design of synthesis filters. The only change is that $2c$ should be replaced by Mc . Readers may consult with [31] for more detail.

4.4 AN ILLUSTRATIVE EXAMPLE

As mentioned in (4.13), the order of the overall transfer function is reduced by a factor of $\frac{1}{M}$. Thus optimization problems for the design of synthesis filter bank is greatly simplified. Moreover in computing inner-outer factorization, and optimal Hankel approximation, only the unstable factor $Q(z)$ is involved. It follows that the computation complexity is related to k only where k is the order of $Q(z)$, i.e., the number of unstable zeros of $\det H_{AC}(z)$. Hence the computation complexity is reduced significantly as well. Because most of the computational effort is on solving Lyapunov equations, the proposed design algorithms in this paper has a computational complexity in the order of $O(k^3)$. Furthermore, our design algorithms involve only numerically reliable algorithms such as Schur decomposition and Lyapunov equations with low computation complexity. The accuracies of the resulting synthesis filter banks can also be ensured. In the following a three-channel filter bank

example is given by using the proposed \mathcal{H}_∞ design algorithm. Although the design algorithm applies to more general multirate filter banks, to simplify the computation, we assume that $H_2(z) = H_0(-z)$. Moreover, by doing so, PR does not apply so that we will observe the quality of the proposed algorithm.

Example 4.1: A three-channel filter bank.

In this example, the analysis filters are linear phase FIR, designed by using the Remez algorithm which can be easily computed with MATLAB command *remez*. $H_0(z)$ is a 17th order lowpass filter with transition band $[\frac{3\pi}{12}, \frac{5\pi}{12}]$, $H_1(z)$ is an 18th order bandpass filter with left transition band $[\frac{3\pi}{12}, \frac{5\pi}{12}]$ and right transition band $[\frac{7\pi}{12}, \frac{9\pi}{12}]$, and $H_2(z)$ is a 17th order high pass filter with transition band $[\frac{7\pi}{12}, \frac{9\pi}{12}]$. The magnitude Bode plots of these three analysis filters are shown in Figure 4.3. From (4.4), $\det H_{AC}(z)$ is computed which has six zeros outside unit circle. Hence PR can not be achieved. A specification of 0.02 for the reconstruction error over all frequency range is required for the design of synthesis filter bank. \mathcal{H}_∞ optimization is applied which yields time delay $\tilde{d} = 7$. This implies total time delay of $d = 3\tilde{d} = 21$. The corresponding $\tilde{G}(z)$ has an order of 6. Balanced truncation is used to reduce the order of $\tilde{G}(z)$ to 4, for which the specification of 0.02 reconstruction error remains true. It follows that $G(z) = \tilde{G}(z^3)$ has an order of $3 \times 4 = 12$, which leads to the three synthesis filters $\hat{F}_0(z)$, $\hat{F}_1(z)$, and $\hat{F}_2(z)$ having order of 47, 46 and 47, respectively. The Bode magnitude plots of reduced synthesis filters $\hat{F}_0(z)$, $\hat{F}_1(z)$, and $\hat{F}_2(z)$ are shown in Figure 4.4. The magnitude frequency response of the associated reconstruction error, a value of 0.0176, is shown in Figure 4.5. According to (4.1), a three channel system has two possible aliasing error components: $A_1(z)$ and $A_2(z)$. Figure 4.6 shows the magnitude frequency response of $A_1(z)$ and $A_2(z)$ which can be regarded as zero.

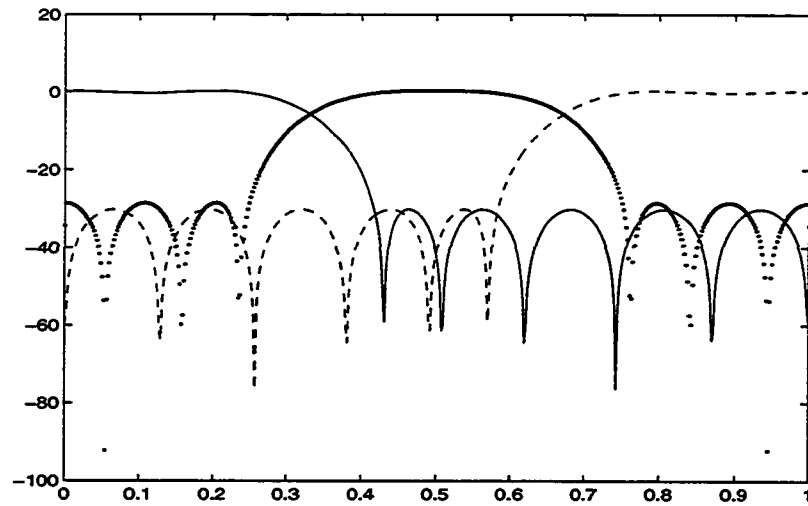


Figure 4.3: Example 4.1: $|H_0|$, $|H_1|$, and $|H_2|$ in dB vs. ω/π .

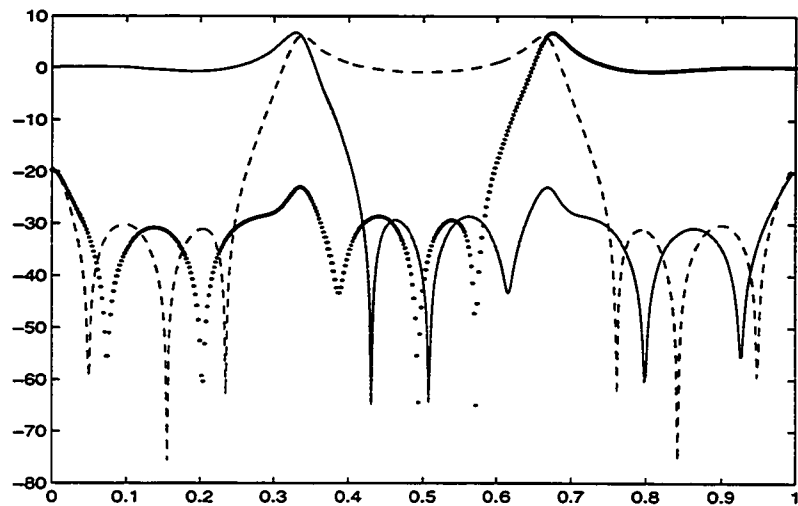


Figure 4.4: Example 4.1: $|\hat{F}_0|$, $|\hat{F}_1|$, and $|\hat{F}_2|$ in dB vs. ω/π .

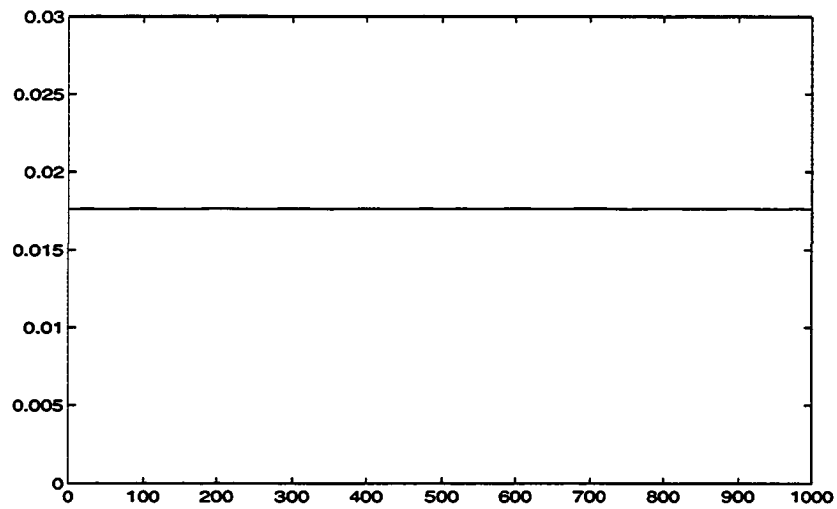


Figure 4.5: Example 4.1: Reconstruction error vs. $1000\omega/\pi$.

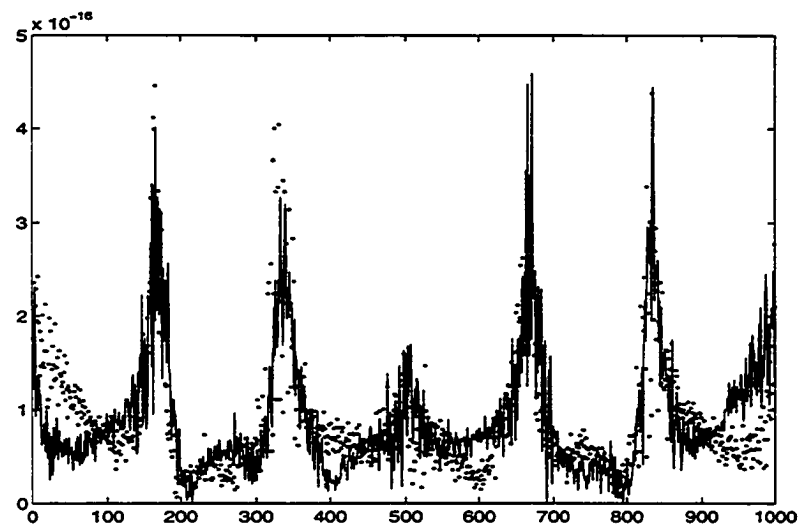


Figure 4.6: Example 4.1: Aliasing error (solid curve: $|A_1|$; dotted curve: $|A_2|$).

4.5 EXTENSION TO NONUNIFORM FILTER BANKS

Nonuniform filter banks are the most general and complicated members in the family of multirate systems. They have far more important application value than that of uniform filter banks. This is because in the real world, nonuniform filter banks give better signal processing result than uniform filter banks in many cases. A general M -channel nonuniform filter bank is shown in Figure 4.7. The integers q_i and p_i are called the decimation and expansion integers, respectively, with reference to the analysis end. Each channel has a sampling factor p_i/q_i . In the analysis bank, the splitted input signals are upsampled, filtered, and downsampled. Through coding and transmitting, the analysis bank makes the combined output a reconstruction of the input signal. One can easily identify Figure 4.7 as a uniform filter bank if $p_i = 1$ and $q_i = M$ for $i = 0, 1, \dots, M - 1$. The setup in Figure 4.7 allows arbitrary (fractional) nonuniform band splitting to be accomplished. For example, by suitable choice of the analysis filters, the frequency band $[0, \pi]$ can be split into the following M subbands:

$$\left[0, \frac{p_0}{q_0}\pi\right), \left[\frac{p_0}{q_0}\pi, \frac{p_0}{q_0}\pi + \frac{p_1}{q_1}\pi\right), \dots, \left[\pi \sum_{i=0}^{M-2} \frac{p_i}{q_i}, \pi \sum_{i=0}^{M-1} \frac{p_i}{q_i}\right).$$

These bands are disjoint when

$$\sum_{i=0}^{M-1} \frac{p_i}{q_i} = 1$$

which also preserves the sampling density.

One direct method for designing perfect reconstruction filter banks with arbitrary rational sampling rate changes was presented in [43]. Unfortunately, this method cannot be used to design some general nonuniform filter banks in which q_i 's are not

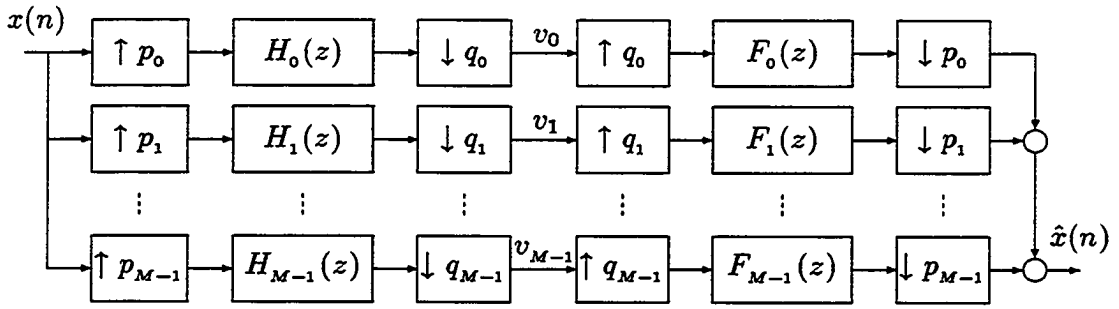


Figure 4.7: The general multirate filter system.

all the same. Specifically if we try to convert this type of nonuniform filter banks into uniform ones, the resulting uniform filter banks will have two or more channels having the same analysis structure. That is, those channels have analysis filters differed from just a scalar factor. This leads to singular aliasing component matrix $H_{AC}(z)$ so that the design of feasible synthesis banks can never be accomplished. If this happens, then the nonuniform filter bank is said to have problem of structural dependency [4, 43]. The problem of structural dependency remains an open research topic for quite a few years, despite great effort from many researchers. On the other hand, one exciting discovery is that a nonuniform filter bank can be transformed into a uniform one directly [4, 43] without structure dependency if q_i 's are all the same with p_i and q_i relatively prime (which is also called coprime). A related work in [59] introduced the idea of generating nonuniform filter banks from uniform filter banks. The procedure of transformation involves polyphase decomposition and noble identities [90]. In particular, formulae were proposed in [43] for the transformation which converts a nonuniform analysis channel into several uniform analysis channels. However, formulae for the corresponding nonuniform synthesis channel were not mentioned. Motivated by the existing work, this section seeks

formulae for the transform in the synthesis bank so that the missing part of [43] will be filled in. Note that some basic but important concepts for the multirate systems and filter banks have already been introduced in Chapter 1. These concepts will be used frequently in this section.

4.5.1 A TWO-CHANNEL NONUNIFORM FILTER BANK

Shown in Figure 4.8 is a two-channel nonuniform multirate filter bank where $q_0 = q_1 = 3$, $p_0 = 2$ and $p_1 = 1$. This typical example has been well known in the literature. Note that this system does not suffer the problem of structure dependency. The frequency range $[0, \pi)$ is split into the low-frequency band $[0, 2\pi/3)$ for channel 0 and the high-frequency band $[2\pi/3, \pi)$ for channel 1. Employing polyphase decomposition yields the expressions:

$$\begin{aligned} H_0(z) &= \sum_{n=-\infty}^{\infty} h(n)z^{-n} = \sum_{n=-\infty}^{\infty} h(2n)z^{-2n} + z^{-1} \sum_{n=-\infty}^{\infty} h(2n+1)z^{-2n} \\ &= E_0(z^2) + z^{-1}E_1(z^2), \end{aligned} \quad (4.20)$$

$$\begin{aligned} F_0(z) &= \sum_{n=-\infty}^{\infty} f(n)z^{-n} = \sum_{n=-\infty}^{\infty} f(2n)z^{-2n} + z^{-1} \sum_{n=-\infty}^{\infty} f(2n+1)z^{-2n} \\ &= R_0(z^2) + z^{-1}R_1(z^2). \end{aligned} \quad (4.21)$$

Channel 0 in Figure 4.8 is actually equivalent to a two-channel uniform system. In the analysis bank, Figure 4.9 shows five equivalent sub-figures. The second and third sub-figures are directly from (4.20). By using the two noble identities shown in Figure 1.3, the fourth sub-figure is obtained. And finally, since 2 and 3 are coprime, according to Figure 1.2, the decimator and the expander can be interchanged so that the fifth sub-figure is valid. Similarly, in the synthesis bank, the equivalent system is shown in Figure 4.10. Therefore, channel 0 in Figure 4.8 is equivalent to cascading

the fifth sub-figure of Figure 4.9 with the fifth sub-figure of Figure 4.10. Furthermore, according to Figure 1.5, Figure 4.11 represents a 2×2 identity matrix which can be easily extended to $M \times M$ case. Combine Figure 4.9, 4-10, and 4-11 together, channel 0 in Figure 4.8 is equivalent to Figure 4.12 so that the whole two-channel nonuniform system in Figure 4.8 becomes a three-channel uniform system shown in Figure 4.13. Note that in the synthesis bank, the procedure of the transformation is the counterpart of that in the analysis bank by using polyphase decomposition. Next we will illustrate the case of multi-channel filters which are not so trivial.

4.5.2 MULTI-CHANNEL NONUNIFORM FILTER BANKS

In a nonuniform filter bank system, if p and q are coprime with $q > p > 1$, then any single channel with decimation ratio $\frac{p}{q}$ can be transformed into a p -channel uniform filter bank. Combining those channels with decimation ratio $\frac{1}{q}$, the nonuniform system can then be totally transformed into a uniform system, which is also called maximally decimated filter banks [90]. For $p \geq 3$, although a polyphase decomposition of $H(z)$ lists its components in order with the index such as $E_0(z)$, $E_1(z)$, $E_2(z)$ etc., they may not appear in the same order in the block configuration. This fact was illustrated in [43]. The same will be true for the corresponding synthesis bank. According to [90], a p -fold polyphase decomposition of $H(z)$ can be written as

$$\begin{aligned}
 H(z) &= \sum_{n=-\infty}^{\infty} h(np)z^{-np} + z^{-1} \sum_{n=-\infty}^{\infty} h(np+1)z^{-np} \\
 &\quad + \cdots + z^{-(p-1)} \sum_{n=-\infty}^{\infty} h(np+(p-1))z^{-np} \\
 &= E_0(z^p) + z^{-1}E_1(z^p) + \cdots + z^{-(p-1)}E_{p-1}(z^p). \tag{4.22}
 \end{aligned}$$

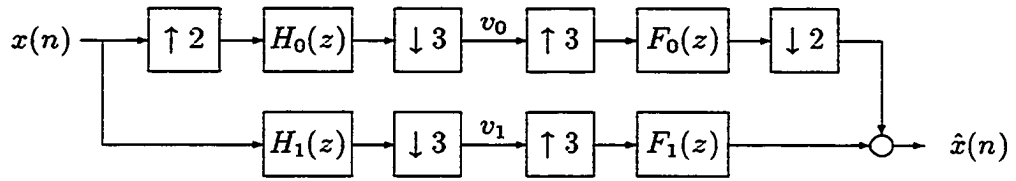


Figure 4.8: A two-channel nonuniform system.

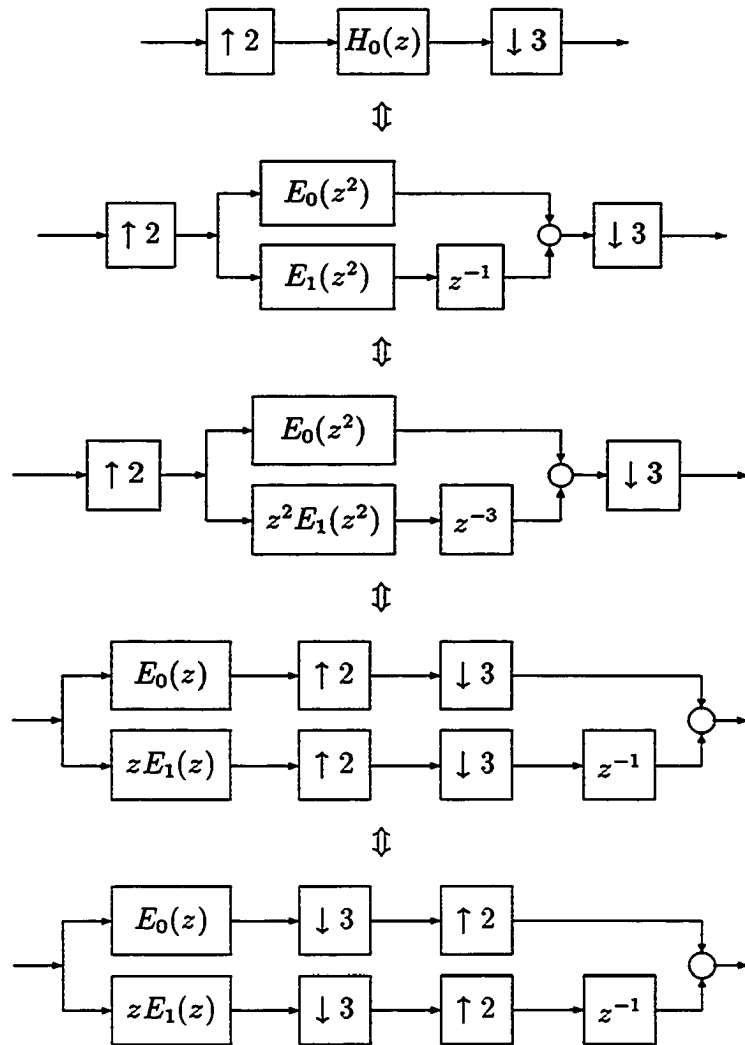


Figure 4.9: Equivalent system of an analysis filter bank branch.

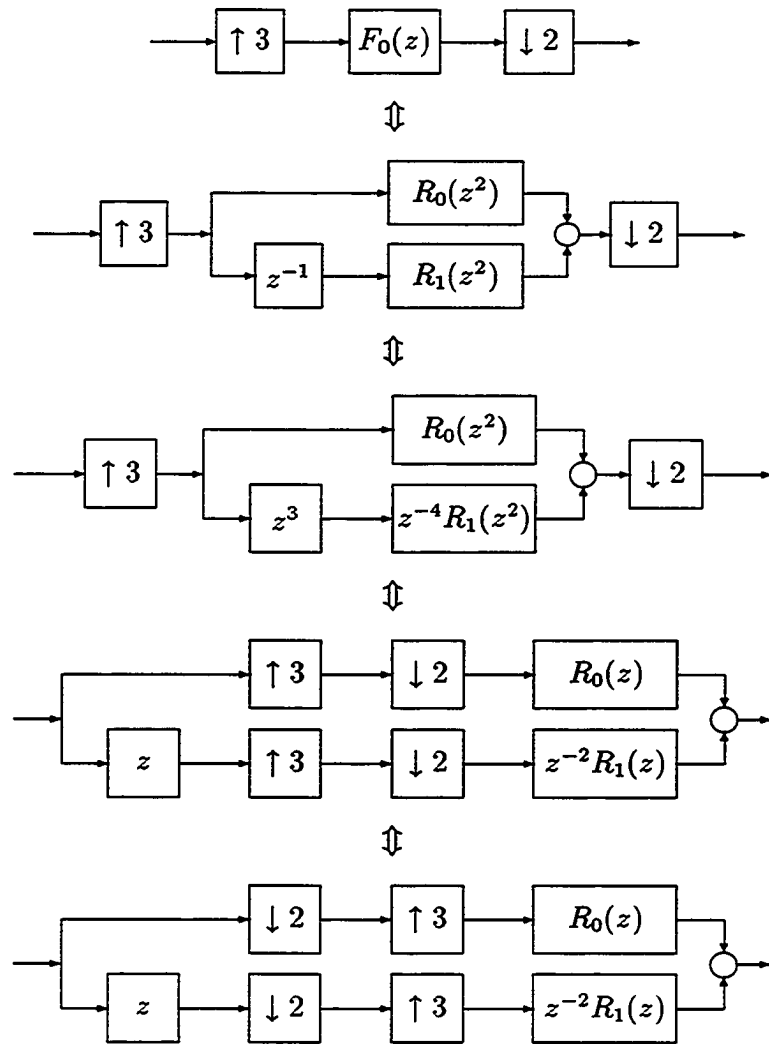


Figure 4.10: Equivalent system of a synthesis filter bank branch.

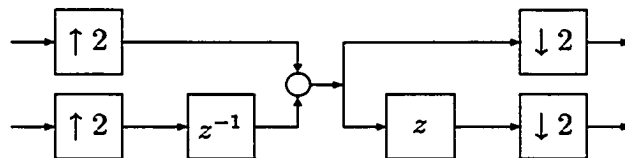


Figure 4.11: An equivalent system of a 2×2 identity matrix.

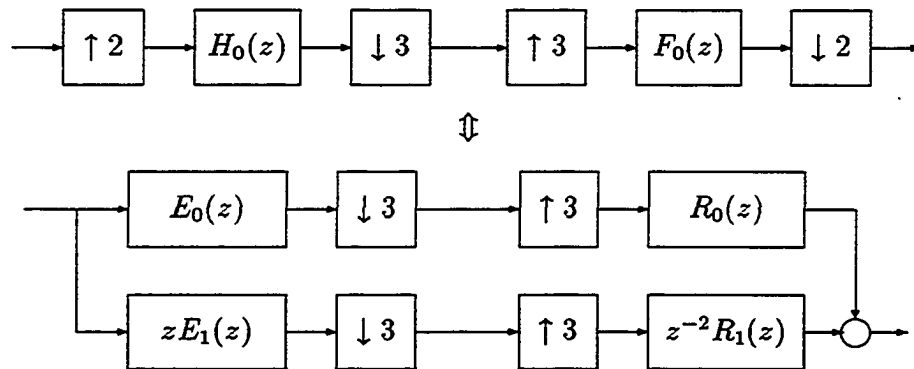


Figure 4.12: Two equivalent systems.

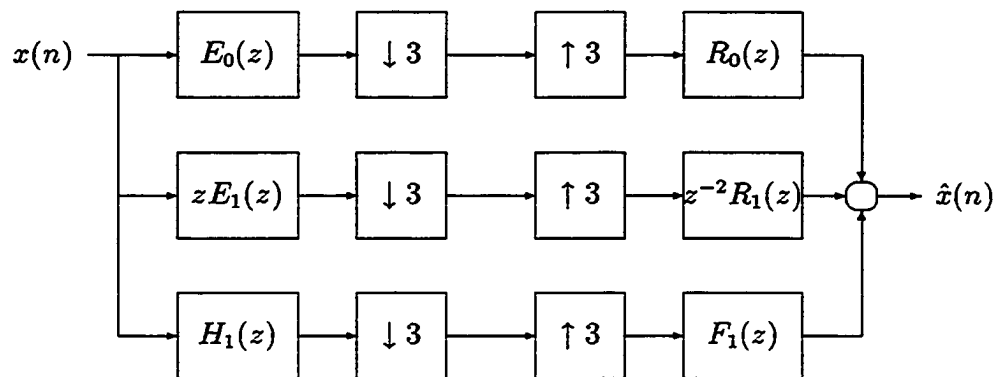


Figure 4.13: An equivalent uniform system to Figure 4.8.

Assume $q_0 = q_1 = \dots = q_{M-1} = q$ with p and q coprime, [43] showed an equivalent system called Transform 1 which is shown here in Figure 4.14 (note $q > p$, and only a single band is shown). For the analysis filters $H'_i(z)$ in Figure 4.14, with $i = 0, 1, \dots, p-1$, [43] gave the following formula,

$$H'_i(z) = z^{d_i} H_{t_i}(z) = z^{d_i} E_{t_i}(z) \quad (4.23)$$

where $d_i = \lfloor \frac{q \cdot i}{p} \rfloor$, $t_i = (q \cdot i) \bmod p$. And $\lfloor x \rfloor$ = the largest integer $\leq x$. Therefore, the nonuniform system can be transformed into a uniform one by using this elegant result. However, since p and q are relatively prime, the formulae for the analysis bank can not be directly used when dealing with the synthesis bank. Because it is difficult to find an expression by using the concept of module, [43] did not give the corresponding explicit formulae for the synthesis bank.

To obtain a desirable uniform filter bank, based on Figure 4.14, the synthesis bank should be in the form of Figure 4.15 (p, q coprime) in which

$$S'_i(z) = z^{-n_i} S_i(z), \quad i = 0, 1, \dots, p-1, \quad (4.24)$$

$$n_0 = 0, \quad n_i p > i \cdot q \quad \text{and} \quad (n_i - 1)p < i \cdot q \quad (4.25)$$

for $i = 1, 2, \dots, p-1$. The conditions in (4.25) requires a unique positive integer n_i for each i . Combine Figure 4.14 and 4.15 with the property shown in Figure 4.11, the whole uniform system is shown in Figure 4.16. From (4.24), Figure 4.17 is derived from Figure 4.15 by using the concepts shown in Figure 1.2 and Figure 1.3. Obviously,

$$F(z) = \sum_{i=0}^{p-1} z^{-n_i \cdot p} S_i(z^p) z^{i \cdot q} = \sum_{i=0}^{p-1} z^{-(n_i \cdot p - i \cdot q)} S_i(z^p) \quad (4.26)$$

with n_i satisfying (4.25). Formula (4.26) shares some similar feature to that in [59] (equation (45)). Given the nonuniform analysis filter bank, the synthesis filter banks can be designed by designing the equivalent uniform system first. A certain delay units may be added to make sure that the final nonuniform synthesis filter bank is causal. However, the difference between our approach and that of [59] is that by imposing (4.25), (4.26) is exactly a p -fold polyphase decomposition of $F(z)$ by the definition in [90] which we will show next. Because [43] gives the formulae (4.23) which lead exactly to the polyphase decomposition in the analysis bank. Our approach matches the idea of [43] and the resulting formulae are the counterpart of the formulae given in (4.23). For the synthesis filter $F(z)$ in Figure 4.17, its polyphase decomposition can be written as

$$\begin{aligned}
 F(z) &= \sum_{n=-\infty}^{\infty} f(np)z^{-np} + z^{-1} \sum_{n=-\infty}^{\infty} f(np+1)z^{-np} \\
 &\quad + \cdots + z^{-(p-1)} \sum_{n=-\infty}^{\infty} f(np+(p-1))z^{-np} \\
 &= R_0(z^p) + z^{-1}R_1(z^p) + \cdots + z^{-(p-1)}R_{p-1}(z^p). \tag{4.27}
 \end{aligned}$$

Lemma 4.1 *Given coprime positive integers p and q with $q > p > 1$. If positive integers n_i satisfy $n_i \cdot p > i \cdot q$ and $(n_i - 1) \cdot p < i \cdot q$ for $i = 1, 2, \dots, p-1$, then $1 \leq n_i \cdot p - i \cdot q \leq p-1$.*

Proof: From the given condition, $0 < n_i \cdot p - i \cdot q < p$. since p, q, i , and n_i are integers, $1 \leq n_i \cdot p - i \cdot q \leq p-1$. ■

Lemma 4.2 *Given coprime positive integers p and q with $q > p > 1$. If positive integers n_i satisfy $n_i \cdot p > i \cdot q$ and $(n_i - 1) \cdot p < i \cdot q$ for $i = 1, 2, \dots, p-1$, then $n_i \cdot p - i \cdot q \neq n_j \cdot p - j \cdot q$ for all $i \neq j$ with $j = 1, 2, \dots, p-1$.*

Proof: Suppose $n_i \cdot p - i \cdot q = n_j \cdot p - j \cdot q$. Without loss of generality, assume $j > i$. This implies $(n_j - n_i) \cdot p = (j - i)q = kq$ with integers $k \in [1, p - 2]$. But since p and q are coprime, $k_{\min} = \frac{(n_j - n_i)p}{q} = p$ which is a contradiction. Therefore, $n_i \cdot p - i \cdot q \neq n_j \cdot p - j \cdot q$ for $i \neq j$. ■

Lemma 4.1 and 4.2 serve as the proof of the following result.

Theorem 4.6 *Given coprime positive integers p and q with $q > p > 1$. If positive integers n_i satisfy $n_i \cdot p > i \cdot q$ and $(n_i - 1) \cdot p < i \cdot q$ for $i = 1, 2, \dots, p - 1$, then with $n_0 = 0$, (4.26) is a p -fold polyphase decomposition of $F(z)$.*

Remark 4.1 *Compare (4.26) with (4.27),*

$$S_0(z) = R_0(z), \quad S_i(z) = R_{n_i p - i q}(z), \quad i = 1, 2, \dots, p - 1. \quad (4.28)$$

Clearly, in (4.27) the polyphase components are ordered while in (4.26) they are not. Mathematically (4.26) and (4.27) are the same. But (4.26) certainly reflects the real world situation.

Corollary 4.1 *Given coprime positive integers p and q with $q > p > 1$. If positive integers n_i satisfy $n_i \cdot p > i \cdot q$ and $(n_i - 1) \cdot p < i \cdot q$ for $i = 1, 2, \dots, p - 1$, then $0 < |(n_{i+1} - n_i)p - q| < p$ for every $(i + 1) \in [1, p - 1]$.*

Proof: By Lemma 4.2, $(n_{i+1} - n_i)p - q \neq 0$. Based on the given conditions, $(i + 1)q < n_{i+1} \cdot p < (i + 1)q + p$ while $-i \cdot q - p < -n_i \cdot p < -i \cdot q$. Add them together,

$$0 < q - p < (n_{i+1} - n_i)p < q + p. \quad (4.29)$$

Therefore, $0 < |(n_{i+1} - n_i)p - q| < p$. ■

Corollary 4.1 actually states that the finite sequence of n_i is monotonically increasing with (4.29) as the proof. Finally, by applying the similar idea to the analysis bank, there exists a polyphase decomposition

$$H(z) = \sum_{i=0}^{p-1} A_i(z^p) z^{(m_i p - i \cdot q)}$$

$$m_0 = 0, \quad m_i p - i \cdot q < 0 \quad \text{and} \quad (m_i + 1)p - i \cdot q > 0. \quad (4.30)$$

Moreover, with $E_i(z)$ defined in (4.22), a virtually same result as (4.23) given by [43] can be written as

$$A'_i(z) = z^{m_i} A_i(z), \quad (4.31)$$

$$A_0(z) = E_0(z), \quad A_i(z) = E_{i \cdot q - m_i p}(z), \quad i = 1, 2, \dots, p-1 \quad (4.32)$$

Figure 4.14 can be redrawn as shown in Figure 4.18. The following corollary is similar to Lemma 4.1 and 4.2 so the proof is omitted.

Corollary 4.2 *Given coprime positive integers p and q with $q > p > 1$. If positive integers m_i satisfy $m_i \cdot p < i \cdot q$ and $(m_i + 1) \cdot p > i \cdot q$ for $i = 1, 2, \dots, p-1$, then $1 \leq -(m_i \cdot p - i \cdot q) \leq p-1$. Moreover, $m_i \cdot p - i \cdot q \neq m_j \cdot p - j \cdot q$ for all $i \neq j$ with $j = 1, 2, \dots, p-1$.*

Example 4.2: Suppose that in Figure 4.14, the decimation ratio $p/q = 4/7$. The nonuniform channel can be converted into a uniform analysis bank with four channels. For $i = 0, 1, 2, 3$, in order to satisfy (4.25) and (4.30) respectively,

$$n_0 = 0, \quad n_1 = 2, \quad n_2 = 4, \quad n_3 = 6; \quad n_i p - i \cdot q = 0, 1, 2, 3,$$

$$m_0 = 0, \quad m_1 = 1, \quad m_2 = 3, \quad m_3 = 5; \quad m_i p - i \cdot q = 0, -3, -2, -1.$$

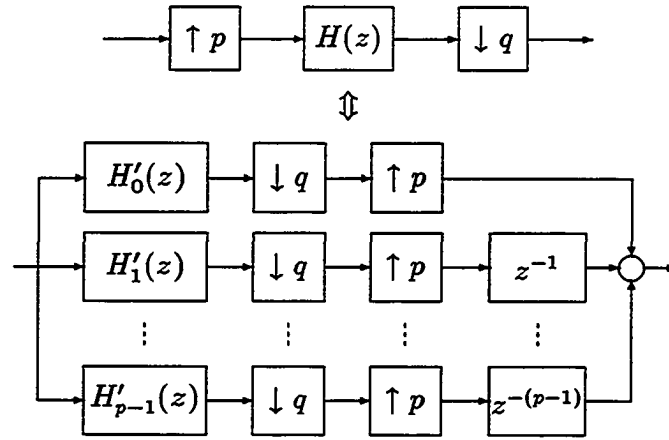
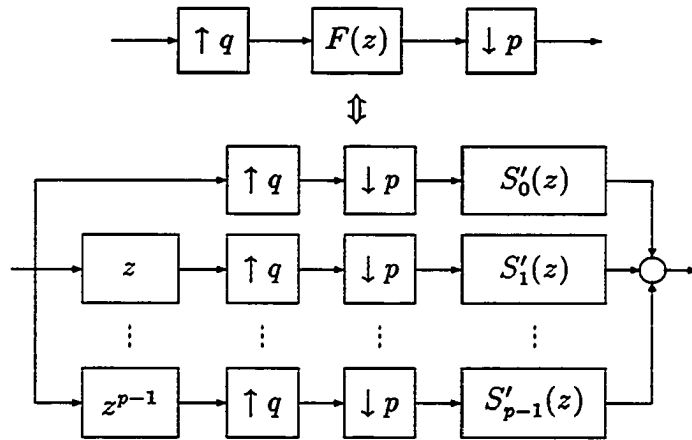


Figure 4.14: Kovacevic's Transform 1.

Figure 4.15: Desired synthesis bank relating to Figure 4-14, $S'_i(z) = z^{-n_i} S_i(z)$.

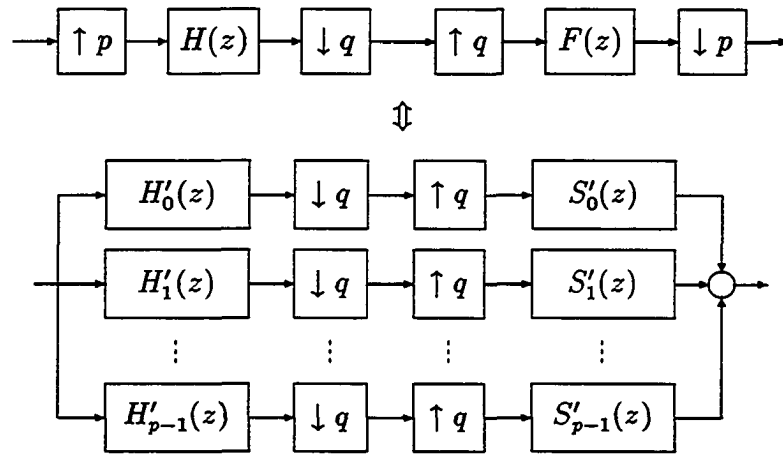


Figure 4.16: Desired uniform system, $S'_i(z) = z^{-n_i} S_i(z)$.

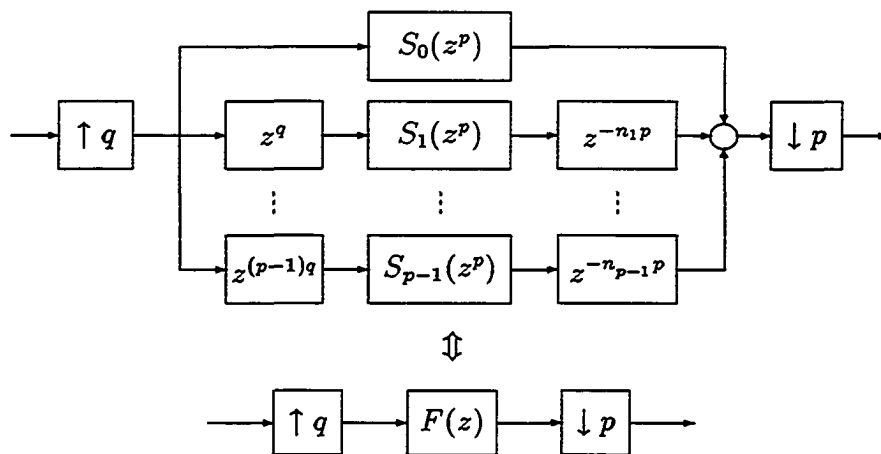


Figure 4.17: Equivalent system to Figure 4.15.

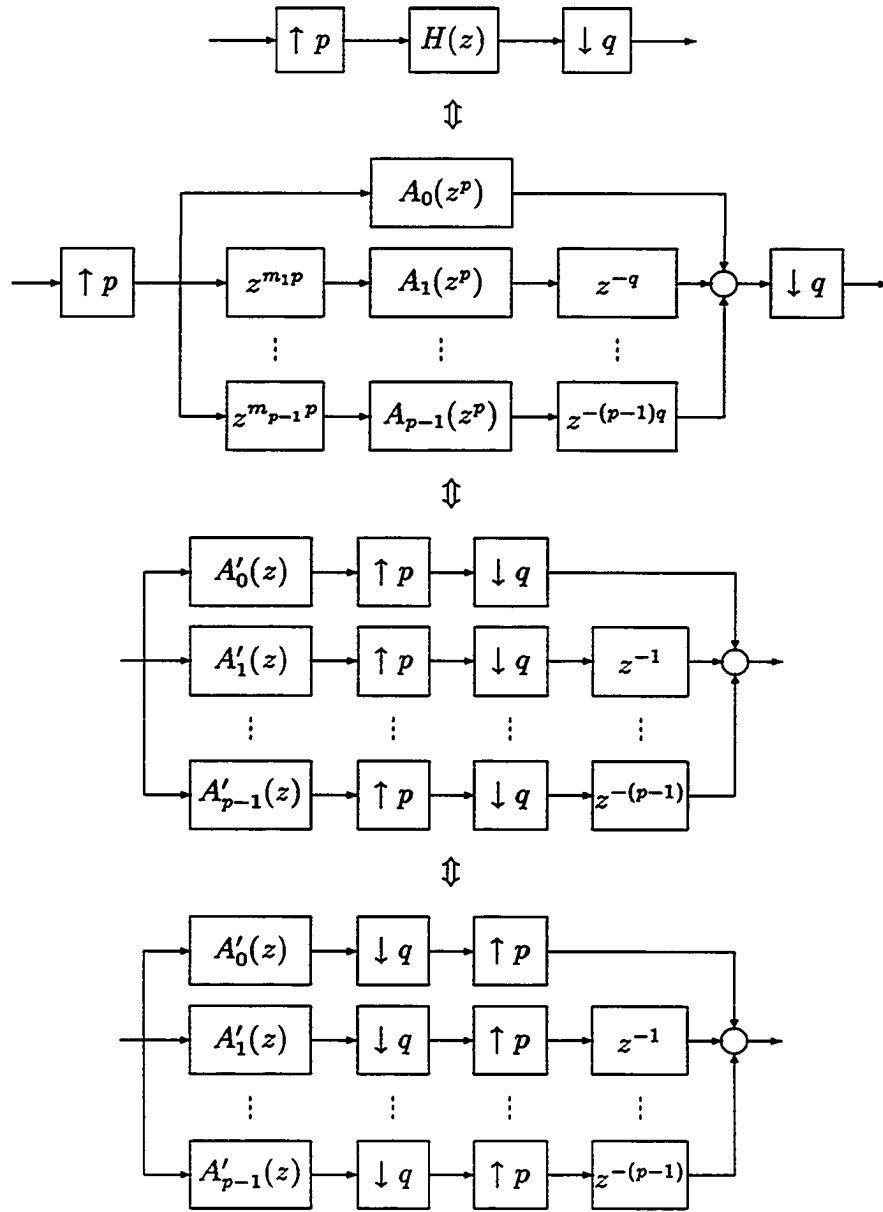


Figure 4.18: Same analysis bank as in Figure 4.14, $A'_i(z) = z^{m_i} A_i(z)$.

For the analysis bank in Figure 4.14, applying (4.23), $H'_i(z) = z^{d_i} E_{t_i}(z)$, which is given in [43],

$$H'_0(z) = E_{t_0}(z) = E_0(z),$$

$$H'_1(z) = z E_{t_1}(z) = z E_3(z),$$

$$H'_2(z) = z^3 E_{t_2}(z) = z^3 E_2(z),$$

$$H'_3(z) = z^5 E_{t_3}(z) = z^5 E_1(z).$$

Now apply (4.31) and (4.32) to Figure 4.18, by no surprise, the result is exactly the same. That is,

$$A'_0(z) = z^{m_0} A_0(z) = E_0(z),$$

$$A'_1(z) = z^{m_1} A_1(z) = z E_3(z),$$

$$A'_2(z) = z^{m_2} A_2(z) = z^3 E_2(z),$$

$$A'_3(z) = z^{m_3} A_3(z) = z^5 E_1(z).$$

On the other hand, for the synthesis bank in Figure 4.15, (4.23) does not work. However, applying (4.24) and (4.28),

$$S'_0(z) = z^{-n_0} S_0(z) = R_0(z),$$

$$S'_1(z) = z^{-n_1} S_1(z) = z^{-2} R_1(z),$$

$$S'_2(z) = z^{-n_2} S_2(z) = z^{-4} R_2(z),$$

$$S'_3(z) = z^{-n_3} S_3(z) = z^{-6} R_3(z).$$

CHAPTER 5

DESIGN OF 2-D QMF BANKS

5.1 INTRODUCTION

Multidimensional (M-D) multirate techniques find applications in subband coding of images and video data [99, 98, 92, 90]. The design of 1-D perfect reconstruction filter banks employing linear phase has been well known. The design of 2-D multirate filter banks, which is extended from the 1-D case, has been reported in [76]. While perfect reconstruction was preferred, it may result in trivial filter banks if each individual filter is restricted to be linear phase with small amplitude distortion. However QMF banks can be designed to achieve alias-free and linear phase with small amplitude distortion (AMD) via optimization. In this chapter, we adopt the Johnston's method to design 2-D two-channel diamond-shaped linear phase filter banks extended from the 1-D case [2, 25, 34]. The design criterion is of least-square type with an error index which is a linear combination of the reconstruction error and the stop band error in quadratic norm [36]. In fact, the generalization from 1-D to 2-D is very challenging. First, the analytical expressions for both the stop band error and the reconstruction error are very difficult to obtain. The derivations are very tedious and long. Second, there is no similar relation between the two analysis filters as in the 1-D case. Each filter has to be designed individually. Third, the iterative algorithm proposed in this chapter, though convergent, is different from the 1-D case. The properties for the 1-D iterative algorithm in [25] may not hold for the 2-D iterative algorithm. The main contribution of this chapter is the explicit

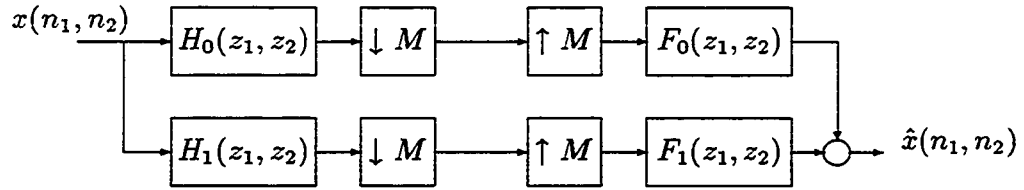


Figure 5.1: A 2-D two-channel multirate filter bank.

expression of the error index, and development of an iterative algorithm for the design of 2-D QMF bank.

5.2 ERROR INDEX FUNCTION AND 2-D LINEAR PHASE FILTERS

The 2-D two-channel multirate filter bank system is shown in Figure 5.1. The analysis filters $H_0(z_1, z_2)$ and $H_1(z_1, z_2)$ are the diamond-shaped lowpass and high-pass analysis filters, respectively, and $F_0(z_1, z_2)$ and $F_1(z_1, z_2)$ are the synthesis filters. Similar to the 1-D case, the input/output relation in \mathcal{Z} -transform for Figure 5.1 is given by

$$\begin{aligned} \hat{X}(z_1, z_2) &= \frac{1}{2} [H_0(z_1, z_2)F_0(z_1, z_2) + H_1(z_1, z_2)F_1(z_1, z_2)] X(z_1, z_2) \\ &+ \frac{1}{2} [H_0(-z_1, -z_2)F_0(z_1, z_2) + H_1(-z_1, -z_2)F_1(z_1, z_2)] X(-z_1, -z_2). \end{aligned} \quad (5.1)$$

The term involving $X(-z_1, -z_2)$ is called aliasing error which can be completely eliminated by choosing $F_0(z_1, z_2) = 2H_1(-z_1, -z_2)$, and $F_1(z_1, z_2) = -2H_0(-z_1, -z_2)$.

Thus

$$\begin{aligned} \hat{X}(z_1, z_2) &= T(z_1, z_2)X(z_1, z_2) \\ &= [H_0(z_1, z_2)H_1(-z_1, -z_2) - H_0(-z_1, -z_2)H_1(z_1, z_2)] X(z_1, z_2). \end{aligned} \quad (5.2)$$

Perfect reconstruction will be achieved if the pure-delay condition is imposed such that

$$T(z_1, z_2) := H_0(z_1, z_2)H_1(-z_1, -z_2) - H_0(-z_1, -z_2)H_1(z_1, z_2) = z_1^{-d_1} z_2^{-d_2}. \quad (5.3)$$

Otherwise, just like in the 1-D case optimization should be applied to minimize the amplitude distortion while keeping zero alias error and linear phase of the system. Note that linear phase is a more important issue in the 2-D case which has been widely discussed in the image processing area.

Throughout this chapter it is assumed that $H_\ell(z_1, z_2)$ is FIR, linear phase, and admits quadrantly symmetric frequency response for both $\ell = 0$ and $\ell = 1$. Hence [12, 76]

$$H_\ell(e^{j\omega_1}, e^{j\omega_2}) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} h_\ell(k_1, k_2) e^{-jk_1\omega_1} e^{-jk_2\omega_2} \quad (5.4)$$

$$= e^{-j\frac{N_1-1}{2}\omega_1} e^{-j\frac{N_2-1}{2}\omega_2} e^{j\frac{M_\ell\pi}{2}} R_\ell(\omega_1, \omega_2) \quad (5.5)$$

for some integer M_ℓ where $h_\ell(k_1, k_2)$ is the impulse response, and $R_\ell(\omega_1, \omega_2)$ is a positive real function of ω_1 and ω_2 . According to [76], the following choices

$$M_0 = 0, \quad N_{01} = 2n_{01} + 1, \quad N_{02} = 2n_{02} + 1, \quad (5.6)$$

$$M_1 = 0, \quad N_{11} = 2n_{11} + 1, \quad N_{12} = 2n_{12} + 1, \quad (5.7)$$

are suitable with n_{ik} some integers such that $n_{01} + n_{02}$ is odd (even) and $n_{11} + n_{12}$ even (odd). Let delay constants $d_1 = n_{11} - n_{01}$ and $d_2 = n_{12} - n_{02}$. Define

$$R(\omega_1, \omega_2) := R_0(\omega_1, \omega_2)R_1(\omega_1 + \pi, \omega_2 + \pi) + R_0(\omega_1 + \pi, \omega_2 + \pi)R_1(\omega_1, \omega_2). \quad (5.8)$$

Then in Figure 5.1 the transfer function between the input and the output can be rewritten as

$$T(e^{j\omega_1}, e^{j\omega_2}) = e^{-j(n_{01}+n_{11})\omega_1} e^{-j(n_{02}+n_{12})\omega_2} R(\omega_1, \omega_2).$$

Clearly, perfect reconstruction requires that $R(\omega_1, \omega_2) = 1$ for all ω_1 and ω_2 . It can be shown that

$$|R(z_1, z_2) - 1|^2 = |G_1^T(z_1^{-1}, z_2^{-1})G_0(z_1, z_2) - z_1^{d_1} z_2^{d_2}|^2 \quad (5.9)$$

where $G_0(z_1, z_2)$ and $G_1(z_1, z_2)$ are defined as

$$G_0(z_1, z_2) = \begin{bmatrix} H_0(-z_1, -z_2) \\ H_0(z_1, z_2) \end{bmatrix}, \quad G_1(z_1, z_2) = \begin{bmatrix} H_1(z_1, z_2) \\ -H_1(-z_1, -z_2) \end{bmatrix}. \quad (5.10)$$

In this chapter, filters $H_0(z_1, z_2)$ and $H_1(z_1, z_2)$ will be designed to minimize the error index function

$$E = (1 - \alpha)E_r + \alpha E_s, \quad 0 < \alpha < 1, \quad (5.11)$$

where E_r is the reconstruction error defined by

$$E_r = \frac{1}{\pi^2} \int_0^\pi \int_0^\pi |T(e^{j\omega_1}, e^{j\omega_2}) - e^{-jd_1\omega_1} e^{-jd_2\omega_2}|^2 d\omega_1 d\omega_2, \quad (5.12)$$

E_s is the stop band energy defined by

$$E_s = \frac{1}{\pi^2} \int_\delta^\pi \int_{\pi-\omega_1+\delta}^\pi R_0^2(\omega_1, \omega_2) d\omega_2 d\omega_1 + \frac{1}{\pi^2} \int_0^{\pi-\delta} \int_0^{\pi-\omega_1-\delta} R_1^2(\omega_1, \omega_2) d\omega_2 d\omega_1, \quad (5.13)$$

due to the diamond shape of the lowpass filter, with δ the frequency of the stop band edge. The minimization of E is well understood for the 1-D case where explicit formula for the expression of E in terms of impulse response of $H_i(z)$ are available. However, its generalization to the 2-D case is much harder due to the evaluation of the stop band error and the reconstruction error.

5.3 COMPUTATION OF THE STOP BAND ERROR E_s

This section focuses on the stop band error E_s . The amplitude response of the analysis filter $H_0(z_1, z_2)$ is now written as

$$R_0(\omega_1, \omega_2) = \sum_{k_1=0}^{\frac{N_1-1}{2}} \sum_{k_2=0}^{\frac{N_2-1}{2}} a(k_1, k_2) \cos(k_1 \omega_1) \cos(k_2 \omega_2). \quad (5.14)$$

The relation between $a(k_1, k_2)$ and $h_0(k_1, k_2)$ is given by

$$\begin{aligned} a(0, 0) &= h_0\left(\frac{N_{01}-1}{2}, \frac{N_{02}-1}{2}\right), \\ a(0, k_2) &= 2h_0\left(\frac{N_{01}-1}{2}, \frac{N_{02}-1}{2} - k_2\right), \quad k_2 = 1, 2, \dots, \frac{N_{02}-1}{2}, \\ a(k_1, 0) &= 2h_0\left(\frac{N_{01}-1}{2} - k_1, \frac{N_{02}-1}{2}\right), \quad k_1 = 1, 2, \dots, \frac{N_{01}-1}{2}, \\ a(k_1, k_2) &= 4h_0\left(\frac{N_{01}-1}{2} - k_1, \frac{N_{02}-1}{2} - k_2\right), \quad k_i = 1, \dots, \frac{N_{0i}-1}{2}, \quad i = 1, 2. \end{aligned}$$

Let A be the coefficient matrix of $R_0(\omega_1, \omega_2)$ defined by

$$A = \begin{bmatrix} a(0, 0) & a(0, 1) & \dots & a(0, \frac{N_{02}-1}{2}) \\ a(1, 0) & a(1, 1) & \dots & a(1, \frac{N_{02}-1}{2}) \\ \vdots & \vdots & \dots & \vdots \\ a(\frac{N_{01}-1}{2}, 0) & a(\frac{N_{01}-1}{2}, 1) & \dots & a(\frac{N_{01}-1}{2}, \frac{N_{02}-1}{2}) \end{bmatrix}. \quad (5.15)$$

By abuse of notation, denote H_0 as the impulse response matrix corresponding to $H_0(z_1, z_2)$ such that

$$\begin{aligned}
 H_0 &= \begin{bmatrix} h_0(0,0) & h_0(0,1) & \cdots & h_0(0, N_{02}-1) \\ h_0(1,0) & h_0(1,1) & \cdots & h_0(1, N_{02}-1) \\ \vdots & \vdots & \cdots & \vdots \\ h_0(N_{01}-1,0) & h_0(N_{01}-1,1) & \cdots & h_0(N_{01}-1, N_{02}-1) \end{bmatrix} \\
 &= \begin{bmatrix} \Omega_1 \\ \frac{1}{4}\Omega_4 \end{bmatrix} A \begin{bmatrix} \Omega_2 & \Omega_3 \end{bmatrix} = \Upsilon_a A \Sigma_a \quad (5.16)
 \end{aligned}$$

$$\begin{aligned}
 \Omega_1 &= \begin{bmatrix} 0_{n_{01} \times 1} & \frac{1}{4}\tilde{I}_{n_{01}} \\ \frac{1}{2} & 0_{1 \times n_{01}} \end{bmatrix}, \quad \Omega_2 = \begin{bmatrix} 0_{1 \times n_{02}} & 2 \\ \tilde{I}_{n_{02}} & 0_{n_{02} \times 1} \end{bmatrix}, \\
 \Omega_3 &= \begin{bmatrix} 0_{1 \times n_{02}} \\ I_{n_{02}} \end{bmatrix}, \quad \Omega_4 = \begin{bmatrix} 0_{n_{01} \times 1} & I_{n_{01}} \end{bmatrix}.
 \end{aligned}$$

Note that matrix \tilde{I}_α is defined in Chapter 2. Rewrite the amplitude response of the analysis filter $H_0(z_1, z_2)$ as

$$\begin{aligned}
 R_0(\omega_1, \omega_2) &= c_1^T(\omega_1) A c_2(\omega_2) = a^T c(\omega_1, \omega_2), \\
 c_k^T(\omega_1) &= \begin{bmatrix} 1 & \cos \omega_k & \cdots & \cos n_{0k} \omega_k \end{bmatrix}, \quad k = 1, 2, \\
 c(\omega_1, \omega_2) &= \begin{bmatrix} 1 \cdot c_2(\omega_2) \\ \cos(\omega_1) \cdot c_2(\omega_2) \\ \vdots \\ \cos(\frac{N_{01}-1}{2}\omega_1) \cdot c_2(\omega_2) \end{bmatrix}, \\
 a^T &= \begin{bmatrix} A_{1,*} & A_{2,*} & \cdots & A_{i,*} & \cdots & A_{n_{01}+1,*} \end{bmatrix},
 \end{aligned}$$

with $A_{i,*}$ the i th row of matrix A . Similarly, $R_1(\omega_1, \omega_2)$ has exactly the same form as $R_0(\omega_1, \omega_2)$ has except for a different size. To distinguish with vector a , we use b to associate with $R_1(\omega_1, \omega_2)$, and the corresponding B matrix has a size of $\frac{N_{11}+1}{2} \times \frac{N_{12}+1}{2}$. Thus regarding the stop band error in (5.13), $E_s = E_{s_1} + E_{s_2}$ with

$$\begin{aligned}
 E_{s_1} &= \frac{1}{\pi^2} \int_{\delta}^{\pi} \int_{\pi-\omega_1+\delta}^{\pi} |H_0(e^{j\omega_1}, e^{j\omega_2})|^2 d\omega_2 d\omega_1 \\
 &= \frac{1}{\pi^2} \int_{\delta}^{\pi} \int_{\pi-\omega_1+\delta}^{\pi} R_0^2(\omega_1, \omega_2) d\omega_2 d\omega_1 \\
 &= \frac{1}{\pi^2} \int_{\delta}^{\pi} \int_{\pi-\omega_1+\delta}^{\pi} (a^T c(\omega_1, \omega_2))(a^T c(\omega_1, \omega_2))^T d\omega_2 d\omega_1 \\
 &= a^T \left(\frac{1}{\pi^2} \int_{\delta}^{\pi} \int_{\pi-\omega_1+\delta}^{\pi} c(\omega_1, \omega_2) c^T(\omega_1, \omega_2) d\omega_2 d\omega_1 \right) a = a^T \Phi^{(0)} a, \quad (5.17)
 \end{aligned}$$

$$\begin{aligned}
 E_{s_2} &= \frac{1}{\pi^2} \int_0^{\pi-\delta} \int_0^{\pi-\omega_1-\delta} |H_1(e^{j\omega_1}, e^{j\omega_2})|^2 d\omega_2 d\omega_1 \\
 &= \frac{1}{\pi^2} \int_0^{\pi-\delta} \int_0^{\pi-\omega_1-\delta} R_1^2(\omega_1, \omega_2) d\omega_2 d\omega_1 \\
 &= \frac{1}{\pi^2} \int_0^{\pi-\delta} \int_0^{\pi-\omega_1-\delta} (b^T c(\omega_1, \omega_2))(b^T c(\omega_1, \omega_2))^T d\omega_2 d\omega_1 \\
 &= b^T \left(\frac{1}{\pi^2} \int_0^{\pi-\delta} \int_0^{\pi-\omega_1-\delta} c(\omega_1, \omega_2) c^T(\omega_1, \omega_2) d\omega_2 d\omega_1 \right) b = b^T \Phi^{(1)} b. \quad (5.18)
 \end{aligned}$$

To evaluate matrix $\Phi^{(0)}$, we partition it into $\frac{N_{01}+1}{2} \times \frac{N_{02}+1}{2}$ block matrix with each one square of size $\frac{N_{02}+1}{2} \times \frac{N_{02}+1}{2}$. Similarly, to evaluate $\Phi^{(1)}$, we partition it into $\frac{N_{11}+1}{2} \times \frac{N_{12}+1}{2}$ block matrix with each one square of size $\frac{N_{12}+1}{2} \times \frac{N_{12}+1}{2}$. Therefore, for $k = 0, 1$, matrices $\Phi^{(0)}$ and $\Phi^{(1)}$ have the form

$$\Phi^{(k)} = \begin{bmatrix} \Phi_{1,1} & \Phi_{1,2} & \cdots & \Phi_{1, \frac{N_{k1}+1}{2}} \\ \Phi_{2,1} & \Phi_{2,2} & \cdots & \Phi_{2, \frac{N_{k1}+1}{2}} \\ \vdots & \vdots & \vdots & \vdots \\ \Phi_{\frac{N_{k1}+1}{2},1} & \Phi_{\frac{N_{k1}+1}{2},2} & \cdots & \Phi_{\frac{N_{k1}+1}{2}, \frac{N_{k1}+1}{2}} \end{bmatrix}.$$

Computation of $\Phi^{(0)}$

It is not difficult to see that

$$\Phi_{i,l}^{(0)}(k, m) = \frac{1}{\pi^2} \int_{\delta}^{\pi} \int_{\pi-\omega_1+\delta}^{\pi} \cos((i-1)\omega_1) \cos((k-1)\omega_2) \\ \cos((l-1)\omega_1) \cos((m-1)\omega_2) d\omega_2 d\omega_1$$

for $i, l = 1, 2, \dots, \frac{N_{01}+1}{2}$ and $k, m = 1, 2, \dots, \frac{N_{02}+1}{2}$. Denote

$$a_{01} = i + l - 2, \quad a_{02} = i - l, \quad a_{03} = k + m - 2, \quad a_{04} = k - m,$$

$$CONST_1 = (\pi - \delta) \sin(a_{03}(\pi + \delta)),$$

$$CONST_2 = (\pi - \delta) \sin(a_{04}(\pi + \delta)),$$

$$y_1 = \begin{cases} CONST_1, & a_{01} = a_{03} \\ \frac{\cos(a_{01}\delta + a_{03}\pi) - \cos(a_{01}\pi + a_{03}\delta)}{a_{01} - a_{03}}, & a_{01} \neq a_{03} \end{cases}$$

$$y_2 = \frac{\cos(a_{01}\pi - a_{03}\delta) - \cos(a_{01}\delta - a_{03}\pi)}{a_{01} + a_{03}},$$

$$y_3 = \begin{cases} CONST_1, & a_{02} = a_{03} \\ \frac{\cos(a_{02}\delta + a_{03}\pi) - \cos(a_{02}\pi + a_{03}\delta)}{a_{02} - a_{03}}, & a_{02} \neq a_{03} \end{cases}$$

$$y_4 = \begin{cases} CONST_1, & a_{02} = -a_{03} \\ \frac{\cos(a_{02}\pi - a_{03}\delta) - \cos(a_{02}\delta - a_{03}\pi)}{a_{02} + a_{03}}, & a_{02} \neq -a_{03} \end{cases}$$

$$x_1 = \begin{cases} CONST_2, & a_{01} = a_{04} \\ \frac{\cos(a_{01}\delta + a_{04}\pi) - \cos(a_{01}\pi + a_{04}\delta)}{a_{01} - a_{04}}, & a_{01} \neq a_{04} \end{cases}$$

$$x_2 = \begin{cases} CONST_2, & a_{01} = -a_{04} \\ \frac{\cos(a_{01}\pi - a_{04}\delta) - \cos(a_{01}\delta - a_{04}\pi)}{a_{01} + a_{04}}, & a_{01} \neq -a_{04} \end{cases}$$

$$x_3 = \begin{cases} CONST_2, & a_{02} = a_{04} \\ \frac{\cos(a_{02}\delta + a_{04}\pi) - \cos(a_{02}\pi + a_{04}\delta)}{a_{02} - a_{04}}, & a_{02} \neq a_{04} \end{cases}$$

$$x_4 = \begin{cases} CONST_2, & a_{02} = -a_{04} \\ \frac{\cos(a_{02}\pi - a_{04}\delta) - \cos(a_{02}\delta - a_{04}\pi)}{a_{02} + a_{04}}, & a_{02} \neq -a_{04} \end{cases}$$

$$\Delta_{1a} = \frac{\cos a_{01}\pi - \cos a_{01}\delta}{a_{01}^2} + \frac{\cos a_{02}\pi - \cos a_{02}\delta}{a_{02}^2},$$

$$\Delta_{1b} = (\pi - \delta)^2, \quad \Delta_{1c} = \frac{\cos a_{01}\pi - \cos a_{01}\delta}{a_{01}^2} + \frac{(\pi - \delta)^2}{2},$$

$$\Delta_{11} = \frac{y_1 + y_2 + y_3 + y_4}{2a_{03}}, \quad \Delta_{12} = 2 \frac{\cos a_{03}\delta - \cos a_{03}\pi}{a_{03}^2}, \quad \Delta_{13} = \frac{y_1 + y_2 + 2y_3}{2a_{03}},$$

$$\Delta_{21} = \frac{x_1 + x_2 + x_3 + x_4}{2a_{04}}, \quad \Delta_{22} = 2 \frac{\cos a_{04}\delta - \cos a_{04}\pi}{a_{04}^2}, \quad \Delta_{23} = \frac{x_1 + x_2 + 2x_3}{2a_{04}}.$$

Then for $i, l = 1, 2, \dots, (N_{01} + 1)/2$ and $i \neq l$, the elements of $\Phi^{(0)}$ are given by:

$$\Phi_{i,l}^{(0)}(k, m) = \begin{cases} \frac{1}{2\pi^2} \Delta_{1a}, & k = m = 1, \\ \frac{1}{4\pi^2} (\Delta_{1a} - \Delta_{11}), & k = m \geq 2, \\ \frac{-1}{4\pi^2} (\Delta_{11} + \Delta_{21}), & k \neq m, \end{cases} \quad (5.19)$$

$$\Phi_{1,1}^{(0)}(k, m) = \begin{cases} \frac{1}{2\pi^2} \Delta_{1b}, & k = m = 1, \\ \frac{1}{4\pi^2} (\Delta_{1b} - \Delta_{12}), & k = m \geq 2, \\ \frac{-1}{4\pi^2} (\Delta_{12} + \Delta_{22}), & k \neq m, \end{cases} \quad (5.20)$$

$$\Phi_{i,i}^{(0)}(k, m) = \begin{cases} \frac{1}{2\pi^2} \Delta_{1c}, & k = m = 1, \\ \frac{1}{4\pi^2} (\Delta_{1c} - \Delta_{13}), & k = m \geq 2, \\ \frac{-1}{4\pi^2} (\Delta_{13} + \Delta_{23}), & k \neq m. \end{cases} \quad (5.21)$$

Computation of $\Phi^{(1)}$

Similar to $\Phi^{(0)}$, for $i, l = 1, 2, \dots, \frac{N_{11}+1}{2}$ and $k, m = 1, 2, \dots, \frac{N_{12}+1}{2}$, we have

$$\begin{aligned} \Phi_{i,l}^{(1)}(k, m) &= \frac{1}{\pi^2} \int_0^{\pi-\delta} \int_0^{\pi-\omega_1-\delta} \cos((i-1)\omega_1) \cos((k-1)\omega_2) \\ &\quad \cos((l-1)\omega_1) \cos((m-1)\omega_2) d\omega_2 d\omega_1 \end{aligned}$$

$$CONST_3 = (\pi - \delta) \sin(a_{03}(\pi - \delta)),$$

$$CONST_4 = (\pi - \delta) \sin(a_{04}(\pi - \delta)),$$

$$w_1 = \begin{cases} CONST_3, & a_{01} = a_{03} \\ \frac{\cos(a_{03}(\pi-\delta)) - \cos(a_{01}(\pi-\delta))}{a_{01} - a_{03}}, & a_{01} \neq a_{03} \end{cases}$$

$$w_2 = \frac{\cos(a_{01}(\pi - \delta)) - \cos(a_{03}(\pi - \delta))}{a_{01} + a_{03}},$$

$$w_3 = \begin{cases} CONST_3, & a_{02} = a_{03} \\ \frac{\cos(a_{03}(\pi-\delta)) - \cos(a_{02}(\pi-\delta))}{a_{02} - a_{03}}, & a_{02} \neq a_{03} \end{cases}$$

$$w_4 = \begin{cases} CONST_3, & a_{02} = -a_{03} \\ \frac{\cos(a_{02}(\pi-\delta)) - \cos(a_{03}(\pi-\delta))}{a_{02} + a_{03}}, & a_{02} \neq -a_{03} \end{cases}$$

$$r_1 = \begin{cases} CONST_4, & a_{01} = a_{04} \\ \frac{\cos(a_{04}(\pi-\delta)) - \cos(a_{01}(\pi-\delta))}{a_{01} - a_{04}}, & a_{01} \neq a_{04} \end{cases}$$

$$r_2 = \begin{cases} \text{CONST}_4, & a_{01} = -a_{04} \\ \frac{\cos(a_{01}(\pi-\delta)) - \cos(a_{04}(\pi-\delta))}{a_{01} + a_{04}}, & a_{01} \neq -a_{04} \end{cases}$$

$$r_3 = \begin{cases} \text{CONST}_4, & a_{02} = a_{04} \\ \frac{\cos(a_{04}(\pi-\delta)) - \cos(a_{02}(\pi-\delta))}{a_{02} - a_{04}}, & a_{02} \neq a_{04} \end{cases}$$

$$r_4 = \begin{cases} \text{CONST}_4, & a_{02} = -a_{04} \\ \frac{\cos(a_{02}(\pi-\delta)) - \cos(a_{04}(\pi-\delta))}{a_{02} + a_{04}}, & a_{02} \neq -a_{04} \end{cases}$$

$$\Gamma_{1a} = \frac{1 - \cos a_{01}(\pi - \delta)}{a_{01}^2} + \frac{1 - \cos a_{02}(\pi - \delta)}{a_{02}^2},$$

$$\Gamma_{1b} = (\pi - \delta)^2, \quad \Gamma_{1c} = \frac{1 - \cos a_{01}(\pi - \delta)}{a_{01}^2} + \frac{(\pi - \delta)^2}{2},$$

$$\Gamma_{11} = \frac{w_1 + w_2 + w_3 + w_4}{2a_{03}}, \quad \Gamma_{12} = 2 \frac{1 - \cos a_{03}(\pi - \delta)}{a_{03}^2}, \quad \Gamma_{13} = \frac{w_1 + w_2 + 2w_3}{2a_{03}},$$

$$\Gamma_{21} = \frac{r_1 + r_2 + r_3 + r_4}{2a_{04}}, \quad \Gamma_{22} = 2 \frac{1 - \cos a_{04}(\pi - \delta)}{a_{04}^2}, \quad \Gamma_{23} = \frac{r_1 + r_2 + 2r_3}{2a_{04}}.$$

Then for $i, l = 1, 2, \dots, (N_{11} + 1)/2$ and $i \neq l$, the elements of $\Phi^{(1)}$ are given by:

$$\Phi_{i,l}^{(1)}(k, m) = \begin{cases} \frac{1}{2\pi^2} \Gamma_{1a}, & k = m = 1, \\ \frac{1}{4\pi^2} (\Gamma_{1a} + \Gamma_{11}), & k = m \geq 2, \\ \frac{1}{4\pi^2} (\Gamma_{11} + \Gamma_{21}), & k \neq m, \end{cases} \quad (5.22)$$

$$\Phi_{1,1}^{(1)}(k, m) = \begin{cases} \frac{1}{2\pi^2} \Gamma_{1b}, & k = m = 1, \\ \frac{1}{4\pi^2} (\Gamma_{1b} + \Gamma_{12}), & k = m \geq 2, \\ \frac{1}{4\pi^2} (\Gamma_{12} + \Gamma_{22}), & k \neq m, \end{cases} \quad (5.23)$$

$$\Phi_{i,i}^{(1)}(k, m) = \begin{cases} \frac{1}{2\pi^2} \Gamma_{1c}, & k = m = 1, \\ \frac{1}{4\pi^2} (\Gamma_{1c} + \Gamma_{13}), & k = m \geq 2, \\ \frac{1}{4\pi^2} (\Gamma_{13} + \Gamma_{23}), & k \neq m. \end{cases} \quad (5.24)$$

5.4 COMPUTATION OF THE RECONSTRUCTION ERROR E_r

The computation of the reconstruction error E_r is much more involved, we adopt an approach based on the realization method as well as the Lyapunov equation in linear system theory. Let

$$L_{01} = N_{01} - 1 = 2n_{01}, \quad L_{02} = N_{02} - 1 = 2n_{02},$$

$$L_{11} = N_{11} - 1 = 2n_{11}, \quad L_{12} = N_{12} - 1 = 2n_{12}.$$

Let $|z_2| = 1$ be a parameter. Denote the transfer functions of both the lowpass filter and the highpass filter as

$$H_0(z_1, z_2) = H_0[z_2](z_1) = \sum_{i=0}^{L_{01}} h_{0,i}[z_2] z_1^{-i}, \quad h_{0,i}[z_2] = \sum_{k=0}^{L_{02}} h_0(i, k) z_2^{-k}, \quad (5.25)$$

$$H_1(z_1, z_2) = H_1[z_2](z_1) = \sum_{i=0}^{L_{11}} h_{1,i}[z_2] z_1^{-i}, \quad h_{1,i}[z_2] = \sum_{k=0}^{L_{12}} h_1(i, k) z_2^{-k}. \quad (5.26)$$

Then simple realizations can be defined for

$$G_0(z_1, z_2) = G_0[z_2](z_1),$$

$$G_1(z_1, z_2) = G_1[z_2](z_1),$$

where G_0 and G_1 are as in (5.10). Indeed, define

$$\begin{aligned}
A_t &= \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}_{L_{t1} \times L_{t1}}, \quad B_t = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{L_{t1} \times 1}, \quad t = 0, 1, \\
C_0[z_2] &= \begin{bmatrix} -h_{0,1}(-z_2) & h_{0,2}(-z_2) & \cdots & (-1)^k h_{0,k}(-z_2) & \cdots & h_{0,L_{01}}(-z_2) \\ h_{0,1}(z_2) & h_{0,2}(z_2) & \cdots & h_{0,k}(z_2) & \cdots & h_{0,L_{01}}(z_2) \end{bmatrix}, \\
C_1[z_2] &= \begin{bmatrix} h_{1,1}(z_2) & h_{1,2}(z_2) & \cdots & h_{1,k}(z_2) & \cdots & h_{1,L_{11}}(z_2) \\ h_{1,1}(-z_2) & -h_{1,2}(-z_2) & \cdots & (-1)^{k+1} h_{1,k}(-z_2) & \cdots & -h_{1,L_{11}}(-z_2) \end{bmatrix}, \\
D_0[z_2] &= \begin{bmatrix} h_{0,0}(-z_2) & h_{0,0}(z_2) \end{bmatrix}^T, \quad D_1[z_2] = \begin{bmatrix} h_{1,0}(z_2) & -h_{1,0}(-z_2) \end{bmatrix}^T.
\end{aligned} \tag{5.27}$$

As $L_{01} = 2n_{01}$ and $L_{11} = 2n_{11}$, we have

$$G_0[z_2](z_1) = D_0[z_2] + C_0[z_2](z_1 I - A_0)^{-1} B_0,$$

$$G_1[z_2](z_1) = D_1[z_2] + C_1[z_2](z_1 I - A_1)^{-1} B_1.$$

Obviously,

$$G_1^T[z_2^{-1}](z_1^{-1}) = D_1^T[z_2^{-1}] + B_1^T(z_1^{-1} I - A_1^T)^{-1} C_1^T[z_2^{-1}].$$

The perfect reconstruction condition is equivalent to that

$$G_1^T[z_2^{-1}](z_1^{-1}) G_0[z_2](z_1) = I[z_2](z_1)$$

with $I(z_1, z_2) = z_1^{d_1} z_2^{d_2}$. Define

$$E_r^1[z_2] = \frac{1}{2\pi j} \oint_{|z_1|=1} \left| G_1^T[z_2^{-1}](z_1^{-1}) G_0[z_2](z_1) - I[z_2](z_1) \right|^2 dz_1.$$

It follows that the reconstruction error can be computed by solving two integrals one by one such as

$$\begin{aligned}
E_r &= \frac{1}{(2\pi j)^2} \oint_{|z_2|=1} \oint_{|z_1|=1} \left| G_1^T(z_1^{-1}, z_2^{-1}) G_0(z_1, z_2) - I(z_1, z_2) \right|^2 dz_1 dz_2 \\
&= \frac{1}{2\pi j} \oint_{|z_2|=1} \left(\frac{1}{2\pi j} \oint_{|z_1|=1} \left| G_1^T[z_2^{-1}](z_1^{-1}) G_0[z_2](z_1) - I[z_2](z_1) \right|^2 dz_1 \right) dz_2 \\
&= \frac{1}{2\pi j} \oint_{|z_2|=1} E_r^1[z_2] dz_2.
\end{aligned} \tag{5.28}$$

Hence our first task is the computation of $E_r^1(z_2)$ with z_2 as a parameter on the unit circle. Computation of the reconstruction error E_r can then be carried out subsequently. For ease of notation, note that $h_{0,i}(z_2)$ and $h_{1,i}(z_2)$ as in (5.25) and (5.26) can be written as

$$h_{0,i}(z_2) = \sum_{k=0}^{L_{02}} \alpha_{i,k} z_2^{-k}, \quad h_{0,i}(-z_2) = \sum_{k=0}^{L_{02}} (-1)^k \alpha_{i,k} z_2^{-k}, \tag{5.29}$$

$$h_{1,i}(z_2) = \sum_{k=0}^{L_{12}} \beta_{i,k} z_2^{-k}, \quad h_{1,i}(-z_2) = \sum_{k=0}^{L_{12}} (-1)^k \beta_{i,k} z_2^{-k} \tag{5.30}$$

where

$$\alpha_{i,k} = h_0(i, k), \quad \beta_{i,k} = h_1(i, k). \tag{5.31}$$

Applying Parseval's Theorem to $E_r^1(z_2)$ with z_2 as a parameter on the unit circle yields

$$E_r^1(z_2) = |u_{d_1}(z_2) - z_2^{d_2}|^2 + \sum_{i=-L_{11}, i \neq d_1}^{L_{01}} |u_i(z_2)|^2 \tag{5.32}$$

where $u_i(z_2)$ for both positive and negative i are given in the following as

$$u_i(z_2) = \sum_{l=0}^{\min\{L_{01}-i, L_{11}\}} (-1)^{l+i} \{u_{ia}(z_2) - u_{ib}(z_2)\}, \tag{5.33}$$

$$\begin{aligned}
u_{ia}(z_2) &= \left(\sum_{k=0}^{L_{02}} (-1)^k \alpha_{l+i,k} z_2^{-k} \right) \left(\sum_{k=0}^{L_{12}} \beta_{l,k} z_2^k \right), \\
u_{ib}(z_2) &= (-1)^i \left(\sum_{k=0}^{L_{12}} (-1)^k \beta_{l,k} z_2^k \right) \left(\sum_{k=0}^{L_{02}} \alpha_{l+i,k} z_2^{-k} \right), \\
i &= 0, 1, \dots, L_{01}, \\
u_{-i}(z_2) &= \sum_{l=0}^{\min\{L_{01}, L_{11}-i\}} (-1)^{l+i-1} \{u_{-ia}(z_2) - u_{-ib}(z_2)\}, \\
u_{-ia}(z_2) &= \left(\sum_{k=0}^{L_{12}} (-1)^k \beta_{l+i,k} z_2^k \right) \left(\sum_{k=0}^{L_{02}} \alpha_{l,k} z_2^{-k} \right), \\
u_{-ib}(z_2) &= (-1)^i \left(\sum_{k=0}^{L_{02}} (-1)^k \alpha_{l,k} z_2^{-k} \right) \left(\sum_{k=0}^{L_{12}} \beta_{l+i,k} z_2^k \right), \\
i &= 0, 1, \dots, L_{11}.
\end{aligned} \tag{5.34}$$

The derivation of the above expressions is shown later. Note that for $i > 0$, $u_i(z_2)$ and $-u_{-i}(z_2^{-1})$ have close resemblance in that the positions of α and β are exchanged with necessary modification of the upper limit in summation. Hence, in order to obtain closed expression of the reconstruction error, we need to evaluate only

$$\frac{1}{2\pi j} \int_0^{2\pi} |u_i(z_2)|^2 dz_2, \quad \text{and} \quad \frac{1}{2\pi j} \int_0^{2\pi} |u_{d_1}(z_2) - z_2^{d_2}|^2 dz_2, \quad i \geq 0.$$

Denote the following matrices as

$$P_l^{(i)}(z_2) = (-1)^l \sum_{k=0}^{L_{12}} \begin{bmatrix} \beta_{l,k} \\ (-1)^k \beta_{l,k} \end{bmatrix} z_2^{-k}, \quad P_i(z_2) = \begin{bmatrix} P_0^{(i)}(z_2) \\ \vdots \\ P_{L_{01}-i}^{(i)}(z_2) \end{bmatrix}, \tag{5.35}$$

$$Q_l^{(i)}(z_2) = \sum_{k=0}^{L_{02}} \begin{bmatrix} (-1)^{k+i} \alpha_{l+i,k} \\ -\alpha_{l+i,k} \end{bmatrix} z_2^{-k}, \quad Q_i(z_2) = \begin{bmatrix} Q_0^{(i)}(z_2) \\ \vdots \\ Q_{L_{01}-i}^{(i)}(z_2) \end{bmatrix}. \tag{5.36}$$

We now focus on how to compute the contour integral of $|v_i(z_2)|^2$ for $i \geq 0$ as efficiently as possible. Without loss of generality, it is assumed that $L_{01} \leq L_{11}$, and thus $\min\{L_{01} - i, L_{11}\} = L_{01} - i$ whenever $i \geq 0$. With the above notations, there holds

$$u_i(z_2) = \sum_{l=0}^{L_{01}-i} [P_l^{(i)}(z_2^{-1})]^T Q_l^{(i)}(z_2) = P_i^T(z_2^{-1}) Q_i(z_2).$$

The realization method and the Lyapunov equations are again used to compute the square-integral of $|u_i(z_2)|$. For this purpose we write $P_i(z_2)$ and $Q_i(z_2)$ using state space realization as

$$P_i(z_2) = D_{p_i} + C_{p_i}(z_2 I - A_p)^{-1} B_p, \quad Q_i(z_2) = D_{q_i} + C_{q_i}(z_2 I - A_q)^{-1} B_q,$$

where A_p, A_q , and B_p, B_q have the same form as in (5.27) except for the different size, and

$$C_{p_i}^{(l)} = (-1)^l \begin{bmatrix} \beta_{l,1} & \beta_{l,2} & \cdots & \beta_{l,k} & \cdots & \beta_{l,L_{12}} \\ -\beta_{l,1} & \beta_{l,2} & \cdots & (-1)^k \beta_{l,k} & \cdots & \beta_{l,L_{12}} \end{bmatrix}, \quad (5.37)$$

$$C_{q_i}^{(l)} = \begin{bmatrix} (-1)^{i+1} \alpha_{l+i,1} & \cdots & (-1)^{i+k} \alpha_{l+i,k} & \cdots & (-1)^i \alpha_{l+i,L_{02}} \\ -\alpha_{l+i,1} & \cdots & -\alpha_{l+i,k} & \cdots & -\alpha_{l+i,L_{02}} \end{bmatrix}, \quad (5.38)$$

$$D_{p_i}^{(l)} = (-1)^l \begin{bmatrix} \beta_{l,0} & \beta_{l,0} \end{bmatrix}^T, \quad D_{q_i}^{(l)} = \begin{bmatrix} (-1)^i \alpha_{l+i,0} & -\alpha_{l+i,0} \end{bmatrix}^T, \quad (5.39)$$

$$C_{p_i} = \begin{bmatrix} C_{p_i}^{(0)} \\ C_{p_i}^{(1)} \\ \vdots \\ C_{p_i}^{(L_{01}-i)} \end{bmatrix}, \quad D_{p_i} = \begin{bmatrix} D_{p_i}^{(0)} \\ D_{p_i}^{(1)} \\ \vdots \\ D_{p_i}^{(L_{01}-i)} \end{bmatrix}, \quad C_{q_i} = \begin{bmatrix} C_{q_i}^{(0)} \\ C_{q_i}^{(1)} \\ \vdots \\ C_{q_i}^{(L_{01}-i)} \end{bmatrix}, \quad D_{q_i} = \begin{bmatrix} D_{q_i}^{(0)} \\ D_{q_i}^{(1)} \\ \vdots \\ D_{q_i}^{(L_{01}-i)} \end{bmatrix}.$$

Let Y_i be the solution to the following Lyapunov equation with its notation defined in the previous equations.

$$Y_i - A_p^T Y_i A_q = C_{p_i}^T C_{q_i} = \sum_{l=0}^{L_{01}-i} [C_{p_i}^{(l)}]^T C_{q_i}^{(l)}. \quad (5.40)$$

Then applying Parseval's Theorem yields (see Section 5.6 for the proof)

$$\begin{aligned} \frac{1}{2\pi j} \oint_{|z_2|=1} |u_i(z_2)|^2 dz_2 &= (D_{p_i}^T C_{q_i} + B_p^T Y_i A_q)(D_{p_i}^T C_{q_i} + B_p^T Y_i A_q)^T \quad (5.41) \\ &+ (C_{p_i}^T D_{q_i} + A_p^T Y_i B_q)^T (C_{p_i}^T D_{q_i} + A_p^T Y_i B_q) + (D_{p_i}^T D_{q_i} + B_p^T Y_i B_q)^2 \\ &= \|D_{p_i}^T C_{q_i} + B_p^T Y_i A_q\|^2 + \|C_{p_i}^T D_{q_i} + A_p^T Y_i B_q\|^2 + |D_{p_i}^T D_{q_i} + B_p^T Y_i B_q|^2. \end{aligned}$$

Through the detailed lengthy and tedious calculation which is shown in Section 5.6, for $i \geq 0$, we obtain

$$|D_{p_i}^T D_{q_i} + B_p^T Y_i B_q|^2 = [(-1)^i - 1]^2 \left\{ \text{tr} [\chi_{L+}^{(i)} B \chi_R^{(0)} A^T] \right\}^2, \quad (5.42)$$

$$\|D_{p_i}^T C_{q_i} + B_p^T Y_i A_q\|^2 = \sum_{t=1}^{L_{02}} [(-1)^{i+t} - 1]^2 \left\{ \text{tr} [\chi_{L+}^{(i)} B \chi_R^{(t)} A^T] \right\}^2, \quad (5.43)$$

$$\|C_{p_i}^T D_{q_i} + A_p^T Y_i B_q\|^2 = \sum_{s=1}^{L_{12}} [(-1)^i - (-1)^s]^2 \left\{ \text{tr} [\chi_{L+}^{(i)} B \chi_R^{(s)} A^T] \right\}^2, \quad (5.44)$$

where first,

$$\chi_{L+}^{(i)} = (S_{N_{01}}^i \Upsilon_a)^T \chi_1 \Upsilon_b, \quad \chi_R^{(0)} = \Sigma_b \chi_2 \Sigma_a^T, \quad \chi_R^{(k)} = \Sigma_b \chi_2 S_{N_{02}}^k \Sigma_a^T.$$

Second, the index k either equals to s or equals to t . And χ_1 and χ_2 are defined later in (5.59) and (5.60) respectively. Third, A is the coefficient matrix of $H_0(z_1, z_2)$ defined in Section 5.3, while B is the corresponding coefficient matrix of $H_1(z_1, z_2)$.

On the other hand, the expressions for the case of negative i are accordingly obtained as

$$|D_{p_i}^T D_{q_i} + B_p^T Y_i B_q|^2 = [1 - (-1)^i]^2 \left\{ \text{tr} [\chi_{L-}^{(i)} B \chi_R^{(0)} A^T] \right\}^2, \quad (5.45)$$

$$\|D_{p_i}^T C_{q_i} + B_p^T Y_i A_q\| = \sum_{t=1}^{L_{02}} [(-1)^t - (-1)^i]^2 \left\{ \text{tr} [\chi_{L-}^{(i)} B \chi_R^{(s)} A^T] \right\}^2, \quad (5.46)$$

$$\|C_{p_i}^T D_{q_i} + A_p^T Y_i B_q\|^2 = \sum_{s=1}^{L_{12}} [1 - (-1)^{i+s}]^2 \left\{ \text{tr} [\chi_{L-}^{(i)} B \chi_R^{(s)} A^T] \right\}^2, \quad (5.47)$$

where $\chi_{L-}^{(i)} = \Upsilon_a^T \chi_1 S_{N_{11}}^i \Upsilon_b$. The reconstruction error can now be computed by summation of (5.42)–(5.44), and (5.45)–(5.47) for different index i .

5.5 AN ITERATIVE ALGORITHM AND ITS CONVERGENCE ANALYSIS

Summarizing the expressions obtained in the previous two sections, we conclude that

$$E = (1 - \alpha)E_r + \alpha E_s = E_q(A, A, B, B) + E_l(A, B) + E_c, \quad (5.48)$$

where E_q is a quadratic function of both A and B , E_l a linear function of A and B , and E_c a constant. Explicit forms of E_q , E_l , and E_c can be easily obtained by using the expressions derived in the previous two sections. Moreover, the error function E is also a function of a^T and b^T , the stacked row vectors of A and B respectively. Therefore, (5.48) can be written as

$$E(a, b) = a^T \Phi_b a - 2a^T \Psi_b + r_b = b^T \Phi_a b - 2b^T \Psi_a + r_a, \quad (5.49)$$

where matrices Φ_b , Φ_a are quadratic functions of vectors b , a , and matrices Ψ_b , Ψ_a are linear functions of b and a , respectively. The constant scalar term r_b involves only

b , and r_a only a . Hence given a (equivalently A), the optimal solution b (equivalently B) can be easily computed, and vice versa. This observation motivates the following iterative algorithm:

Iterative Design Algorithm:

- Step 1: Given initial data b , do the following:
- Step 2: Compute the optimal solution of $E(a, b)$ with b fixed which is given by $a = \Phi_b^{-1}\Psi_b$.
- Step 3: For the solution a obtained in Step 2, compute optimal solution of $E(a, b)$ with a fixed which is given by $b = \Phi_a^{-1}\Psi_a$.
- Step 4: If $\|E(a, b)\| \leq \epsilon$, then STOP. Otherwise, go to Step 2.

End.

Theorem 5.1 *Consider the design problem for 2-D QMF banks in Section 5.2. The proposed iterative design algorithm for minimization of the error criterion in (5.11) is convergent, and it converges to the local optimal solution of the 2-D analysis and synthesis filters as shown in Figure 5.1.*

Proof: Let b_0 be the initial data of b . The iterative algorithm generates a sequence of a_k and b_k for $k = 1, 2, \dots$, which are ordered according to the computation steps. By the property of the minimization in Step 2 and Step 3, there holds

$$E(a_{k+1}, b_{k+1}) \leq E(a_{k+1}, b_k) \leq E(a_k, b_k) \leq \dots \leq E(a_1, b_1). \quad (5.50)$$

Thus the error sequence $E(a_k, b_k)$ decreases monotonically. Since $E(a, b) > 0$ by the positivity of the squared-integral error, the error sequence $E(a_k, b_k)$ admits a

unique limit E_* . Now the question is whether a_k and b_k converge. Suppose that $E(a_{k+1}, b_{k+1}) = E(a_k, b_k)$ for some k . If $E(a_{k+1}, b_k)$ admits an optimal solution, then according to (5.50), $E(a_{k+1}, b_k) = E(a_k, b_k)$ implies that $a_{k+1} = a_k$ as b_k is fixed. Similarly, $E(a_{k+1}, b_{k+1}) = E(a_{k+1}, b_k)$ implies that $b_{k+1} = b_k$ as $a_{k+1} = a_k$ being fixed. Therefore, as $E(a_k, b_k)$ converges to the limit E_* , a_k and b_k converge to a_* and b_* , respectively. On the other hand, suppose that $E(a_{k+1}, b_{k+1}) < E(a_k, b_k)$ for all k . Since E_* exists, there exist some a_{*1} and b_{*1} such that $E_* = E(a_{*1}, b_{*1})$. Through iteration, there is a_{*2} which minimizes $E(a, b_{*1})$ with b_{*1} fixed. Then we have $E_* = E(a_{*1}, b_{*1}) = E(a_{*2}, b_{*1})$ which means $a_{*1} = a_{*2}$. By similar arguments, there exists b_{*2} such that $b_{*1} = b_{*2}$. This concludes the convergence of the iterative algorithm. Since $E(a, b)$ may admit more than one minimum solution, the limit is a local minimum. But each iteration process admits only one limit. ■

5.6 COMPUTATION OF INTEGRAL ERROR

This section gives more detailed derivation to compute the integral error using Lyapunov method. Specifically the quantity to be computed has the form $J_k = \frac{1}{2j\pi} \oint_{|z|=1} |G_1^T(z^{-1})G_0(z)|^\ell dz$, $\ell = 1, 2$, with both $G_0(z)$ and $G_1(z)$ causal and stable. Let (A_0, B_0, C_0, D_0) and (A_1, B_1, C_1, D_1) be the realizations of $G_0(z)$ and $G_1(z)$ respectively. Then the Lyapunov equation $X - A_1^H X A_0 = C_1^H C_0$ has a unique solution X which is positive definite where superscript H denotes conjugate transpose. This leads to the expansion [25]:

$$\begin{aligned} G_1^T(z_1^{-1})G_0(z_1) &= (D_1^H + B_1^H(z^{-1}I - A_1^H)^{-1}C_1^H)(D_0 + C_0(zI - A_0)^{-1}B_0) \\ &= D_1^H D_0 + B_1^H X B_0 + (D_1^H C_0 + B_1^H X A_0)(zI - A_0)^{-1}B_0 \\ &\quad + B_1^T(z^{-1}I - A_1^H)^{-1}(C_1^T(z^{-1})D_0 + A_1^H X B_0). \end{aligned}$$

For $\ell = 1$, the integral error has an expression of $J_1 = D_1^H D_0 + B_1^H X B_0$. For $\ell = 2$, the expression of integral error J_2 is more complicated. However if both $G_0(z)$ and $G_1(z)$ are FIR, then Parseval's Theorem can be applied to compute J_2 .

Proof of (5.32) – (5.34):

The realization of G_0 and G_1 are now functions of z_2 on the unit circle. The Lyapunov equation

$$X - A_1^H X A_0 = C_1^H C_0$$

again has a unique solution $X_2(z_2)$ that is positive definite for all z_2 on the unit circle. The simple realizations of $G_0[z_2](z_1)$ and $G_1[z_2](z_1)$ imply that [37]

$$(z_1 I - A_i)^{-1} B_i = \frac{[z_1^{L_{i1}-1} \ \cdots \ z_1 \ 1]^T}{z_1^{L_{i1}}}.$$

Denote $u_0(z_2) = D_1^T(z_2^{-1})D_0(z_2) + B_1^T X(z_2)B_0$, and

$$\begin{aligned} \begin{bmatrix} u_1(z_2) & u_2(z_2) & \cdots & u_{L_{01}}(z_2) \end{bmatrix} &= D_1^T(z_2^{-1})C_0(z_2) + B_1^T X(z_2)A_0, \\ \begin{bmatrix} u_{-1}(z_2) & u_{-2}(z_2) & \cdots & u_{-L_{11}}(z_2) \end{bmatrix}^T &= (C_1^T(z_2^{-1})D_0(z_2) + A_1^T X(z_2)B_0)^T. \end{aligned}$$

Then there holds

$$G_1^T[z_2^{-1}](z_1^{-1})G_0[z_2](z_1) - I[z_2](z_1) = u_0(z_2) + \sum_{i=1}^{L_{01}} u_i(z_2)z_1^{-i} + \sum_{i=1}^{L_{11}} u_{-i}(z_2^{-1})z_1^{-i} - I(z_2)z_1^{d_1}.$$

Applying Parseval's Theorem to $E_r^1(z_2)$ with z_2 as a parameter on the unit circle yields

$$E_r^1(z_2) = |u_{d_1}(z_2) - z_2^{d_2}|^2 + \sum_{i=-L_{11}, i \neq d_1}^{L_{01}} |u_i(z_2)|^2 \quad (5.51)$$

which is shown in (5.32). Denote $h_{0j}(-z_2) = g_{0j}(z_2)$ for $j = 0, 1, \dots, L_{01}$, and denote $h_{1i}(-z_2) = g_{1i}(z_2)$ for $i = 0, 1, \dots, L_{11}$. Since L_{i1} are even (as N_{i1} are odd) for $i = 0, 1$,

$$C_0(z_2) = \begin{bmatrix} -g_{0,1}(z_2) & \cdots & (-1)^j g_{0,j}(z_2) & \cdots & g_{0,L_{01}}(z_2) \\ h_{0,1}(z_2) & \cdots & h_{0,j}(z_2) & \cdots & h_{0,L_{01}}(z_2) \end{bmatrix},$$

$$C_1(z_2) = \begin{bmatrix} h_{1,1}(z_2) & \cdots & h_{1,i}(z_2) & \cdots & h_{1,L_{11}}(z_2) \\ g_{1,1}(z_2) & \cdots & (-1)^{i-1} g_{1,i}(z_2) & \cdots & -g_{1,L_{11}}(z_2) \end{bmatrix},$$

$$D_0(z_2) = \begin{bmatrix} g_{0,0}(z_2) \\ h_{0,0}(z_2) \end{bmatrix}, \quad D_1(z_2) = \begin{bmatrix} h_{1,0}(z_2) \\ -g_{1,0}(z_2) \end{bmatrix}.$$

Denote $x_{i,j}(z_2)$ as the i - j th element of $X(z_2)$. It can be clearly verified through direct computation [22] that

$$x_{i,L_{01}} = g_{0,L_{01}}(z_2) \bar{h}_{1,i}(z_2) + (-1)^{i-1} \bar{g}_{1,i}(z_2) h_{0,L_{01}}(z_2), \quad \text{for } 1 \leq i \leq L_{11}, \quad (5.52)$$

$$x_{L_{11},j} = (-1)^j g_{0,j}(z_2) \bar{h}_{1,L_{11}}(z_2) - \bar{g}_{1,L_{11}}(z_2) h_{0,j}(z_2), \quad \text{for } 1 \leq j \leq L_{01}, \quad (5.53)$$

where $\bar{h}(z_2) = h(z_2^{-1})$ and $\bar{g}(z_2) = g(z_2^{-1})$ by $|z_2| = 1$. Moreover,

$$\begin{aligned} x_{i,j} &= x_{i+1,j+1} + (-1)^j g_{0,j}(z_2) \bar{h}_{1,i}(z_2) + (-1)^{i-1} \bar{g}_{1,i}(z_2) h_{0,j}(z_2) \\ &= \sum_{k=0}^{\min(L_{01}-j, L_{11}-i)} \{(-1)^{j+k} g_{0,j+k}(z_2) \bar{h}_{1,i+k}(z_2) + (-1)^{i-1+k} \bar{g}_{1,i+k}(z_2) h_{0,j+k}(z_2)\}. \end{aligned} \quad (5.54)$$

Without loss of generality, assume that $L_{11} > L_{01}$, then in particular,

$$x_{1,j} = \sum_{k=0}^{L_{01}-j} \{(-1)^{j+k} g_{0,j+k}(z_2) \bar{h}_{1,1+k}(z_2) + (-1)^k \bar{g}_{1,1+k}(z_2) h_{0,j+k}(z_2)\},$$

$$x_{i,1} = \sum_{k=0}^{\min(L_{01}-1, L_{11}-i)} \{(-1)^{1+k} g_{0,1+k}(z_2) \bar{h}_{1,i+k}(z_2) + (-1)^{i-1+k} \bar{g}_{1,i+k}(z_2) h_{0,1+k}(z_2)\}.$$

Therefore, we obtain the following expressions:

$$D_1^T(z_2^{-1})D_0(z_2) = g_{0,0}(z_2)\bar{h}_{1,0}(z_2) - \bar{g}_{1,0}(z_2)h_{0,0}(z_2),$$

$$B_1^T X(z_2)B_0 = x_{1,1}(z_2) = \sum_{l=1}^{L_{01}} \{(-1)^l g_{0,l}(z_2) \bar{h}_{1,l}(z_2) + (-1)^{l-1} \bar{g}_{1,l}(z_2) h_{0,l}(z_2)\},$$

$$D_1^T(z_2^{-1})D_0(z_2) + B_1^T X(z_2)B_0 = \sum_{l=0}^{L_{01}} \{(-1)^l g_{0,l}(z_2) \bar{h}_{1,l}(z_2) + (-1)^{l-1} \bar{g}_{1,l}(z_2) h_{0,l}(z_2)\}.$$

$$\begin{aligned} D_1^T(z_2^{-1})C_0(z_2) &= \begin{bmatrix} \cdots & (-1)^j g_{0,j}(z_2) \bar{h}_{1,0}(z_2) - \bar{g}_{1,0}(z_2) h_{0,j}(z_2) & \cdots \end{bmatrix}_{L_{01}}, \\ B_1^T X(z_2)A_0 &= \begin{bmatrix} x_{1,2}(z_2) & x_{1,3}(z_2) & \cdots & x_{1,L_{01}}(z_2) & 0 \end{bmatrix}, \\ x_{1,j+1}(z_2) &= \sum_{l=1}^{L_{01}-j} \{(-1)^{j+l} g_{0,j+l}(z_2) \bar{h}_{1,l}(z_2) + (-1)^{l-1} \bar{g}_{1,l}(z_2) h_{0,j+l}(z_2)\}. \end{aligned}$$

$$\begin{aligned} &D_1^T(z_2^{-1})C_0(z_2) + B_1^T X(z_2)A_0 \\ &= \begin{bmatrix} \cdots & \sum_{l=0}^{L_{01}-j} \{(-1)^{j+l} g_{0,j+l}(z_2) \bar{h}_{1,l}(z_2) + (-1)^{l-1} \bar{g}_{1,l}(z_2) h_{0,j+l}(z_2)\} & \cdots \end{bmatrix}_{L_{01}}. \end{aligned}$$

$$[C_1^T(z_2^{-1})D_0(z_2)]^T = \begin{bmatrix} \cdots & g_{0,0}(z_2) \bar{h}_{1,i}(z_2) + (-1)^{i-1} \bar{g}_{1,i}(z_2) h_{0,0}(z_2) & \cdots \end{bmatrix}_{L_{11}},$$

$$(A_1^T X(z_2)B_0)^T = \begin{bmatrix} x_{2,1}(z_2) & x_{3,1}(z_2) & \cdots & x_{L_{11},1}(z_2) & 0 \end{bmatrix}^T,$$

$$x_{i+1,1} = \sum_{l=1}^{\min(L_{01}, L_{11}-i)} (-1)^l g_{0,l}(z_2) \bar{h}_{1,i+l}(z_2) + (-1)^{i+l-1} \bar{g}_{1,i+l}(z_2) h_{0,l}(z_2).$$

$$\begin{aligned} &[C_1^T(z_2^{-1})D_0(z_2) + A_1^T X(z_2)B_0]^T \\ &= \begin{bmatrix} \cdots & \sum_{l=0}^{\min(L_{01}, L_{11}-i)} (-1)^l g_{0,l}(z_2) \bar{h}_{1,i+l}(z_2) + (-1)^{i+l-1} \bar{g}_{1,i+l}(z_2) h_{0,l}(z_2) & \cdots \end{bmatrix}_{L_{11}}. \end{aligned}$$

Substitute (5.29)–(5.31) into the above equations, the expressions in equations (5.33) and (5.34) are obtained. ■

Proof of (5.42) – (5.47):

Denote $y_{s,t}$ as the s - t th elements of Y_i that solves (5.40). Then again by [22],

$$y_{L_{12},t} = \sum_{l=0}^{L_{01}-i} (-1)^l \left[(-1)^{i+t} - 1 \right] \beta_{l,L_{12}} \alpha_{l+i,t}, \quad 1 \leq t \leq L_{02}, \quad (5.55)$$

$$y_{s,L_{02}} = \sum_{l=0}^{L_{01}-i} (-1)^l \left[(-1)^i - (-1)^s \right] \beta_{l,s} \alpha_{l+i,L_{02}}, \quad 1 \leq s \leq L_{12}, \quad (5.56)$$

and generally, for $1 \leq s \leq L_{12} - 1$, $1 \leq t \leq L_{02} - 1$,

$$\begin{aligned} y_{s,t} &= y_{s+1,t+1} + \sum_{l=0}^{L_{01}-i} (-1)^l \left[(-1)^{i+t} - (-1)^s \right] \beta_{l,s} \alpha_{l+i,t} \\ &= \sum_{k=0}^{\min(L_{02}-t, L_{12}-s)} \sum_{l=0}^{L_{01}-i} (-1)^l \left[(-1)^{i+t+k} - (-1)^{s+k} \right] \beta_{l,s+k} \alpha_{l+i,t+k}. \end{aligned} \quad (5.57)$$

According to (5.39),

$$\begin{aligned} D_{p_i}^T D_{q_i} &= \sum_{l=0}^{L_{01}-i} (-1)^l \left[(-1)^i - 1 \right] \beta_{l,0} \alpha_{l+i,0}, \\ B_{p_i}^T Y_i B_{q_i} &= y_{1,1} = \sum_{m=1}^{L_{02}} \sum_{l=0}^{L_{01}-i} (-1)^{l+m} \left[(-1)^i - 1 \right] \beta_{l,m} \alpha_{l+i,m}, \end{aligned}$$

Adding together yields

$$\begin{aligned} D_{p_i}^T D_{q_i} + B_{p_i}^T Y_i B_{q_i} &= \sum_{m=0}^{L_{02}} \sum_{l=0}^{L_{01}-i} (-1)^{l+m} \left[(-1)^i - 1 \right] \beta_{l,m} \alpha_{l+i,m} \\ &= \left[(-1)^i - 1 \right] \text{tr} \left[\chi_1 H_1 \chi_2 (S_{N_{01}}^i H_0)^T \right] \end{aligned} \quad (5.58)$$

where χ_1 and χ_2 are defined respectively as

$$\chi_1 = \text{diag} [(-1)^0, (-1)^1, \dots, (-1)^{N_{01}}] \begin{bmatrix} I_{N_{01}} & 0_{N_{01}, N_{11}-N_{01}} \end{bmatrix}, \quad (5.59)$$

$$\chi_2^T = \text{diag} [(-1)^0, (-1)^1, \dots, (-1)^{N_{02}}] \begin{bmatrix} I_{N_{02}} & 0_{N_{02}, (N_{12}-N_{02})}^T \end{bmatrix}. \quad (5.60)$$

Similarly

$$\begin{aligned} D_{p_i}^T C_{q_i} &= \sum_{l=0}^{L_{01}-i} [D_{p_i}^{(l)}]^T C_{q_i}^{(l)} = \left[\dots \sum_{l=0}^{L_{01}-i} (-1)^l [(-1)^{i+t} - 1] \beta_{l,0} \alpha_{l+i,t} \dots \right]_{L_{02}}, \\ B_p^T Y_i A_q &= \begin{bmatrix} y_{1,2} & y_{1,3} & \dots & y_{1,L_{02}} & 0 \end{bmatrix}, \\ y_{1,t+1} &= \sum_{m=1}^{L_{02}-t} \sum_{l=0}^{L_{01}-i} (-1)^{l+m} [(-1)^{i+t} - 1] \beta_{l,m} \alpha_{l+i,t+m}, \quad t = 1, 2, \dots, L_{02} - 1. \end{aligned}$$

Adding together for $t = 1, 2, \dots, L_{02}$ gives,

$$\begin{aligned} D_{p_i}^T C_{q_i} + B_p^T Y_i A_q &= \begin{bmatrix} \dots \sum_{m=0}^{L_{02}-t} \sum_{l=0}^{L_{01}-i} (-1)^{l+m} [(-1)^{i+t} - 1] \beta_{l,m} \alpha_{l+i,t+m} \dots \end{bmatrix} \\ &= \begin{bmatrix} \dots [(-1)^{i+t} - 1] \text{tr} [\chi_1 H_1 \chi_2 S_{N_{02}}^t H_0^T (S_{N_{01}}^i)^T] \dots \end{bmatrix}. \end{aligned} \quad (5.61)$$

Direct calculation also gives

$$\begin{aligned} (C_{p_i}^T D_{q_i})^T &= \left(\sum_{l=0}^{L_{01}-i} [C_{p_i}^{(l)}]^T D_{q_i}^{(l)} \right)^T \\ &= \begin{bmatrix} \dots \sum_{l=0}^{L_{01}-i} (-1)^l [(-1)^i - (-1)^s] \beta_{l,s} \alpha_{l+i,0} \dots \end{bmatrix}_{L_{12}}, \\ (A_p^T Y_i B_q)^T &= \begin{bmatrix} y_{2,1} & y_{3,1} & \dots & y_{L_{12},1} & 0 \end{bmatrix}, \\ y_{s+1,1} &= \sum_{m=1}^{\min(L_{02}, L_{12}-s)} \sum_{l=0}^{L_{01}-i} (-1)^{l+m} [(-1)^i - (-1)^s] \beta_{l,s+m} \alpha_{l+i,m}, \end{aligned}$$

for $s = 1, 2, \dots, L_{12} - 1$. Adding together yields

$$\begin{aligned}
& (C_{p_i}^T D_{q_i} + A_p^T Y_i B_q)^T \\
&= \begin{bmatrix} \dots & \sum_{m=0}^{\min(L_{02}, L_{12}-s)} \sum_{l=0}^{L_{01}-i} (-1)^{l+m} [(-1)^i - (-1)^s] \beta_{l,s+m} \alpha_{l+i,m} & \dots \end{bmatrix} \\
&= \begin{bmatrix} \dots & [(-1)^i - (-1)^s] \text{tr} [\chi_1 H_1 (S_{N_{12}}^s)^T \chi_2 (S_{N_{01}}^i H_0)^T] & \dots \end{bmatrix}. \tag{5.62}
\end{aligned}$$

Therefore, we obtain

$$\begin{aligned}
|D_{p_i}^T D_{q_i} + B_p^T Y_i B_q|^2 &= [(-1)^i - 1]^2 \left\{ \text{tr} [\chi_1 H_1 \chi_2 (S_{N_{01}}^i H_0)^T] \right\}^2, \\
\|D_{p_i}^T C_{q_i} + B_p^T Y_i A_q\| &= \sum_{t=1}^{L_{02}} [(-1)^{i+t} - 1]^2 \left\{ \text{tr} [\chi_1 H_1 \chi_2 S_{N_{02}}^t (S_{N_{01}}^i H_0)^T] \right\}^2, \\
\|C_{p_i}^T D_{q_i} + A_p^T Y_i B_q\|^2 &= \sum_{s=1}^{L_{12}} [(-1)^i - (-1)^s]^2 \left\{ \text{tr} [\chi_1 H_1 (S_{N_{12}}^s)^T \chi_2 (S_{N_{01}}^i H_0)^T] \right\}^2,
\end{aligned}$$

which can be written equivalently into (5.42)–(5.44). If the index i is negative, then we have

$$\begin{aligned}
|D_{p_i}^T D_{q_i} + B_p^T Y_i B_q|^2 &= [1 - (-1)^i]^2 \left\{ \text{tr} [\chi_1 S_{N_{11}}^i H_1 \chi_2 H_0^T] \right\}^2, \\
\|D_{p_i}^T C_{q_i} + B_p^T Y_i A_q\| &= \sum_{t=1}^{L_{02}} [(-1)^t - (-1)^i]^2 \left\{ \text{tr} [\chi_1 S_{N_{11}}^i H_1 \chi_2 S_{N_{02}}^t H_0^T] \right\}^2, \\
\|C_{p_i}^T D_{q_i} + A_p^T Y_i B_q\|^2 &= \sum_{s=1}^{L_{12}} [1 - (-1)^{i+s}]^2 \left\{ \text{tr} [\chi_1 S_{N_{11}}^i H_1 (S_{N_{12}}^s)^T \chi_2 H_0^T] \right\}^2,
\end{aligned}$$

which can be written equivalently into (5.45)–(5.47). ■

CHAPTER 6

CONCLUSIONS

The design issue of multirate filter banks has been discussed by adopting a system approach. Two iterative algorithms proposed in [2, 34] for the design of QMF banks were examined in Chapter 2. An explicit expression for the error function was derived by using the Lyapunov stability theory. Two new iterative algorithms were developed and their complete convergence results were obtained. Compared with those algorithms in [2, 34], the new algorithms are applicable to a more general nonlinear optimization problem.

The design methods of maximally decimated (or critically sampled) filter banks were developed in both Chapter 3 and Chapter 4. Nonuniform filter banks were considered to be converted into uniform ones. Formulas missing in the literature are given. Frequency domain techniques such as \mathcal{H}_2 and \mathcal{H}_∞ optimization were used to the corresponding design issue. Efficient algorithms were derived using state-space tools that can be implemented with MATLAB. The advantage of those design methods lies in two aspects. First, one can design the analysis filters with regard to the coding requirement of the input signal, and then design the synthesis filters to minimize the reconstruction error for which aliasing error is completely eliminated. That is, one does not have to design all analysis and synthesis filters simultaneously. Second, the proposed design technique uses many well known results, such as optimal Hankel approximation, discrete Lyapunov equation and Riccati equation, minimal realization, and model reduction, from control field that greatly reduces the design difficulty. Furthermore, in Chapter 4, the conversion from nonuniform systems

into uniform ones were studied so that the proposed design method also applies to nonuniform filter banks provided that structural dependency does not exist.

The design of two-channel 2-D QMF banks was presented in Chapter 5. This is an extension of the work in Chapter 2. That is, the design criterion is again of least-square type with an error index which is a linear combination of the reconstruction error and the stop band error in quadratic norm. The results included derivations for an explicit expression of the error index to be minimized, and a convergent iterative algorithm to obtain the minimal solution. A distinguished feature in derivation of the reconstruction error is the use of realization and Lyapunov methods from the linear system theory.

The research work reported in this dissertation is fairly complete. But on the other hand, some issues remain which will be summarized next for future research.

1. The global convergence issue for the two iterative algorithms in Chapter 2.

It should be clear that the minimum solutions for the nonlinear optimization as in Chapter 2 are local in general, although all the design examples in our simulation give very decent results. Whether or not local minimums are close to the global minimum is worth further investigation. A much harder problem is clearly the search for global minimum solution.

2. The perfect reconstruction issue for M -channel filter banks.

The design technique developed in Chapters 3 and 4 yields efficient computation procedures, and low time delay in transmission of the signal. However, it does not achieve perfect reconstruction. Further research is needed to extend the design technique in this dissertation to achieve perfect reconstruction.

3. The structural dependency issue for the design of nonuniform filter banks.

This problem has remained open for almost ten years. This dissertation has developed the new design technique in Chapters 3 and 4, but fails to solve the problem of structural dependency. Future research should be devoted to the design of nonuniform filter banks that resolves the problem of structural dependency.

4. Design issue for 2-D filter banks.

Chapter 5 has generalized the design methods in Chapter 2 to 2-D systems. But extension of the design technique in Chapters 3 and 4 to 2-D filter banks was not investigated in this dissertation which is left for future research.

It is believed that problems in Items 2 and 4 are solvable, even though they are not easy. But the problem in Item 1 is more difficult which may not lead to any solid result due to the lack of corresponding tools for general nonlinear optimization. The most important and difficult problem is Item 3. It is believed that future major development for multirate digital signal processing is hinged to solutions for structural dependency.

BIBLIOGRAPHY

- [1] L. Andrew, V.T. Franques, and V.K. Jain, "Eigen design of quadrature mirror filters", *IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing*, Vol.44, No.9, pp.754-757, September 1997.
- [2] C. Chen and J. Lee, "Design of quadrature mirror filters with linear Phase in the frequency domain", *IEEE Transactions on Circuits and Systems - II: Analog and Digital Signal Processing*, Vol.39, No.9, pp.593-605, September 1992.
- [3] T.-W. Chen and B.A. Francis, "Design of multirate filter banks by \mathcal{H}_∞ optimization", *IEEE Transactions on Signal Processing*, Vol.43, No.12, pp.2822-2830, December 1995.
- [4] T.-W. Chen, "Nonuniform multirate filter banks: analysis and design with an \mathcal{H}_∞ performance measure", *IEEE Transactions on Signal Processing*, Vol.45, No.3, pp.572-582, March 1997.
- [5] T. Chen and P.P. Vaidyanathan, "The role of integer matrices in multidimensional multirate systems", *IEEE Transactions on Signal Processing*, Vol.41, pp.1035-1047, March 1993.
- [6] T. Chen and P.P. Vaidyanathan, "Recent developments in multidimensional multirate systems", *IEEE Transactions on Circuits and Systems for Video Technology* Vol.3, No.2, pp.116-137, April 1993.
- [7] T. Chen and P.P. Vaidyanathan, "Multidimensional multirate filters and filter banks derived from one-dimensional filters", *IEEE Transactions on Signal Processing*, Vol.41, pp.1749-1765, May 1993.
- [8] T. Chen and P.P. Vaidyanathan, "Vector space framework for unification of one- and multidimensional filter bank theory", *IEEE Transactions on Signal Processing*, Vol.42, No.8, pp.2006-2021, August 1994.
- [9] A. Croisier, D. Esteban, and C. Galand, "Perfect channel splitting by use of interpolation/decimation/tree decomposition techniques", *International Conference on Information Sciences and Systems*, pp. 443-446, Patras, Greece, August 1976.
- [10] R.E. Crochiere, "Digital signal processor: sub-band coding," *Bell System Technical Journal*, Vol. 60, pp. 1633-1653, 1981.
- [11] I. Daubechies, *Ten Lectures on Wavelets*, Philadelphia, Pennsylvania: SIAM Press, 1992.

- [12] D.E. Dudgeon and R.M. Mersereau, *Multidimensional digital signal processing*, Englewood Cliffs, NJ: Prentice-Hall, 1984.
- [13] M.M. Ekanayake and K. Premaratne, "Two-channel IIR QMF banks with approximately linear-phase analysis and synthesis filters", *IEEE Transactions on Signal Processing*, Vol.43, No.10, pp.2313-2322, October 1995.
- [14] D. Esteban and C. Galand, "Application of quadrature mirror filter to split-band voice coding schemes, in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 191-195, 1977.
- [15] B.L. Evans, R.H. Bamberger, and J.H. McClellan, "Rules for multidimensional multirate structures", *IEEE Transactions on Signal Processing*, Vol.41, pp.762-771, April 1994.
- [16] B.L. Evans, "Designing commutative cascades of multidimensional upsamplers and downsamplers", *IEEE Signal Processing Letters*, Vol.4, No.11, pp.313-316, November 1997.
- [17] N.J. Fliege, *Multirate Digital Signal Processing*, Chichester, UK: Wiley, 1994.
- [18] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their \mathcal{L}^∞ -error bounds", *International Journal of Control*, Vol.39, No.6, pp.1115-1193, 1984.
- [19] G.H. Golub and C.F. Van Loan, *Matrix Computations*, Baltimore: Johns Hopkins University Press, c1983.
- [20] R.A. Gopinath and C.S. Burrus, "On upsampling, downsampling, and rational sampling rate filter banks", *IEEE Transactions on Signal Processing*, Vol.42, No.4, pp.812-824, April 1994.
- [21] F. Grenez, "Design of quadrature mirror filters by linear programming," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 191-195, 1977.
- [22] G. Gu, P.P. Khargonekar and E. B. Lee, "Approximation of infinite dimensional systems", *IEEE Transactions on Automatic Control*, vol. AC-34, pp. 610-618, June 1989.
- [23] G. Gu and B.A. Shenoi, "A novel approach to the synthesis of recursive digital filters with linear phase", *IEEE Transactions on Circuits and Systems*, Vol.38, No.6, pp.602-612, June 1991.
- [24] G. Gu and J. Huang, "Design of QMF banks and nonlinear optimization", *The 28th Southeastern Symposium on System Theory* (Baton Rouge, Louisiana), 1996, pp.88-92.

- [25] G. Gu and J. Huang, "Convergence results on the design of QMF banks", *IEEE Transactions on Signal Processing*, Vol.46, No.3, pp.758-761, March 1998.
- [26] P.-Q. Hoang and P.P. Vaidyanathan, "Non-uniform multirate filter banks: theory and design", *Proceedings of IEEE International Symposium on Circuits and Systems* (Portland, Oregon), 1989, pp.371-374.
- [27] B.-R. Horng, H. Samueli, and A.N. Willson, Jr., "The design of low-complexity linear-phase FIR filter banks using powers-of-two coefficients with an application to subband image coding", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.1, No.4, pp.318-324, December 1991.
- [28] B.-R. Horng and A.N. Willson, Jr., "Lagrange multiplier approaches to the design of two-channel perfect reconstruction linear phase FIR filter banks", *IEEE Transactions on Signal Processing*, Vol.40, pp.364-374, 1992.
- [29] C.-C. Hsiao, "Polyphase filter matrix for rational sampling rate conversion", *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Dallas, TX, April 1987, pp.2173-2176.
- [30] J. Huang and G. Gu, "A direct approach to the design of QMF banks via frequency domain optimizations", to appear in *IEEE Transactions on Signal Processing*, August 1998.
- [31] J. Huang, G. Gu, and B.A. Shenoi, "Design of multichannel QMF banks via frequency domain optimizations", submitted to *IEEE Transactions on Circuits and Systems II*, 1998.
- [32] J. Huang and G. Gu, "Iterative algorithms for the design of 2-D linear-phase QMF banks", submitted to *IEEE Transactions on Circuits and Systems II*, 1998.
- [33] M. Ikehara and K. Inose, "Design of 2-dimensional linear phase perfect reconstruction FIR filter bank using Lagrange multiplier method", *1996 IEEE ICASSP*, Vol.3, pp.1566-1569, 1996.
- [34] V.K. Jain and R.E. Crochiere, "Quadrature mirror filter design in the time domain", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-32, pp. 353-361, 1984.
- [35] S. Jayasimha and C.G. Hiremath, "Pseudo-QMF banks with near-equiripple performance", *IEEE Transactions on Signal Processing*, Vol. 46, No.1, pp.209-214, January 1998.
- [36] J.D. Johnston, "A filter family designed for use in quadrature mirror filter banks", *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp.291-294, April 1980.

- [37] T. Kailath, *Linear Systems*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1980.
- [38] T.A.C.M. Kaller and I.A. Shah, "A group theoretic approach to multidimensional filter banks: theory and applications", *IEEE Transactions on Signal Processing*, Vol. 44, No.6, pp.1392-1405, June 1996.
- [39] G. Karlsson and M. Vetterli, "Theory of two dimensional multirate filter banks", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol.38, pp.925-937, June 1990.
- [40] E. Kofidis, S. Theodoridis, and N. Kalouptsidis, "On the perfect reconstruction problem in N -band multirate maximally decimated FIR filter banks", *IEEE Transactions on Signal Processing*, Vol.44, No.10, pp.2439-2455, October 1996.
- [41] J. Kovacevic and M. Vetterli, "The commutativity of up/downsampling in multiple dimensions", *IEEE Transactions on Information Theory*, Vol.37, pp.695-698, May 1991.
- [42] J. Kovacevic and M. Vetterli, "Design of multidimensional non-separable regular filter banks and wavelets", *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp.389, 1992.
- [43] J. Kovacevic and M. Vetterli, "Perfect reconstruction filter banks with rational sampling factors", *IEEE Transactions on Signal Processing*, Vol.41, No.6, pp.2047-2066, June 1993.
- [44] R. Lagadec, D. Pelloni, and D. Weiss, "A 2-channel, 16-bit digital sampling rate converter for professional digital audio", *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp.93-96, April 1982.
- [45] J.-H. Lee and S.-C. Huang, "Design of two-channel nonuniform-division maximally decimated filter banks using \mathcal{L}_1 error criteria", *IEE Proceeding: vision, image, and signal processing*, Vol.143, No.2, pp.79-87, April 1996.
- [46] J. Li, T.Q. Nguyen, and S. Tantaratana, "A simple design method for near-perfect-reconstruction nonuniform filter banks", *IEEE Transactions on Signal Processing*, Vol.45, No.8, pp.2105-2109, August 1997.
- [47] Y.C. Lim, R.H. Yang, and S.N. Koh, "The design of weighted minmax quadrature mirror filters", *IEEE Transactions on Signal Processing*, Vol.41, No.5, pp.1780-1788, May 1993.
- [48] Y. Lin and P.P. Vaidyanathan, "Theory and design of two-dimensional filter banks: a review", *Multidimensional Systems and Signal Processing*, Vol.7, pp.263-330, July 1996.

- [49] B. Liu and L.T. Bruton, "The design of nonuniform-band maximally-decimated filter banks", *Proceedings of IEEE International Symposium on Circuits and Systems*, pp.375-378, 1993.
- [50] W.-S. Lu and A. Antoniou, *Two Dimensional Digital Filters*, Marcel Dekker, Inc., 1992.
- [51] R. Manduchi, "Some properties of generalized factorable 2-D FIR filters", *IEEE Transactions on Signal Processing*, Vol.44, No.6, pp.1534-1546, June 1996.
- [52] Y. Meyer, *Wavelets: Algorithms and Applications*, Philadelphia: SIAM, 1993.
- [53] F. Mintzer, "Filters for distortion-free two-band multirate filter banks", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol.33, No.3, pp. 626-630, June 1985.
- [54] M. Minoux, *Mathematical Programming – Theory and Algorithms*, John Wiley and Sons, 1986.
- [55] S.K. Mitra, *Digital Signal Processing: A Computer-Based Approach*, The McGraw-Hill Companies, Inc., 1998.
- [56] T. Nagai, T. Fuchie, and M. Ikehara, "Design of linear phase M -channel perfect reconstruction FIR filter banks", *IEEE Transactions on Signal Processing*, Vol.45, No.9, pp.2380-2387, September 1997.
- [57] K. Nayebi, T.P. Barnwell, and M.J.T. Smith, "The design of perfect reconstruction nonuniform band filter banks", in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Toronto, Canada, pp.1781-1784, May 1991.
- [58] K. Nayebi, T.P. Barnwell, and M.J.T. Smith, "Time domain filter bank analysis: a new design theory", *IEEE Transactions on Signal Processing*, Vol.40, No.6, pp.1412-1429, June 1992.
- [59] K. Nayebi, T.P. Barnwell, and M.J.T. Smith, "Nonuniform filter banks: a reconstruction and design theory", *IEEE Transactions on Signal Processing*, Vol.41, No.3, pp.1114-1127, March 1993.
- [60] K. Nayebi, T.P. Barnwell, and M.J.T. Smith, "Low delay FIR filter banks: design and evaluation", *IEEE Transactions on Signal Processing*, Vol.42, pp.24-31, January 1994.
- [61] K. Nayebi, T.P. Barnwell, and M.J.T. Smith, "On the design of FIR analysis-synthesis filter banks with high computational efficiency", *IEEE Transactions on Signal Processing*, Vol.42, pp.825-834, April 1994.

- [62] T.Q. Nguyen and P.P. Vaidyanathan, "Two-channel perfect reconstruction FIR QMF structures which yield linear phase analysis and synthesis filters", *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol.37, No.5, pp.676-690, May 1989.
- [63] T.Q. Nguyen and P.P. Vaidyanathan, "Structures for M -channel perfect reconstruction FIR QMF banks which yield linear phase analysis filters", *IEEE Transactions on Acoustics, Speech and Signal Processing*, pp.433-446, March 1990.
- [64] T.Q. Nguyen, "Partial spectrum reconstruction using digital filter bank", *IEEE Transactions on Signal Processing*, Vol.41, No.9, pp.2778-2795, September 1993.
- [65] T.Q. Nguyen, "Near-perfect-reconstruction pseudo-QMF banks", *IEEE Transactions on Signal Processing*, Vol.42, pp.65-76, 1994.
- [66] A.V. Oppenheim and R.W. Schaffer, *Digital Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [67] T.W. Parks and C.S. Burrus, *Digital Filter Design*, Wiley-Interscience, 1987.
- [68] A. Petraglie and S.K. Mitra, "High speed A/D conversion incorporating a QMF bank", *IEEE Transactions on Instrumentation and Measurement*, Vol.41, pp.427-431, June 1992.
- [69] S. Phoong, C.W. Kim, P.P. Vaidyanathan, and R. Ansari, "A new class of two-channel biorthogonal filter banks and wavelet bases", *IEEE Transactions on Signal Processing*, Vol.43, No.3, pp.649-665, March 1995.
- [70] G. Pirani and V. Zingarelli, "An analytical formula for the design of quadrature mirror filters", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-32, pp.645-648, June 1984.
- [71] L.R. Rabiner and R.W. Schaffer, *Digital Processing of Speech Signals*, Prentice-Hall, 1978.
- [72] P. Saghizadeh and A.N. Willson, Jr., "A generic approach to the design of M -channel uniform-band perfect-reconstruction linear-phase FIR filter banks", *1995 IEEE ICASSP*, Vol.2, pp.1300-1303, 1995.
- [73] E. Salari and S. Lin, "Low complexity sub-band image coding with pseudo QMF and pyramidal lattice VQ", *IEE Proceedings: vision, image, and signal processing*, Vol.144, No.2, pp.81-88, April 1997.

- [74] R.G. Shenoy, D. Burnside, and T.W Parks, "Linear periodic systems and multirate filter design", *IEEE Transactions on Signal Processing*, Vol. 42, pp.2242-2256, 1994.
- [75] H. Shu, T. Chen, and B.A. Francis, "Minimax design of hybrid multirate filter banks", *IEEE Transactions on Circuits and Systems II*, Vol.44, No.2, pp.120-128, February 1997.
- [76] J. Shyu, "Design of Two-Channel Perfect-Reconstruction Linear-Phase Filter Banks for Subband Image Coding by the Lagrange Multiplier Approach", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.5, No.1, pp.48-51, February 1995.
- [77] M.J.T. Smith and T.P. Barnwell,III, "A procedure for designing exact reconstruction filter banks for tree structured sub-band coders", *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, San Diego, California, March 1984.
- [78] M.J.T. Smith and T.P. Barnwell,III, "Exact reconstruction techniques for tree structured subband coders", *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol.ASSP-34, pp.434-441, June 1986.
- [79] M.J.T. Smith and S.L. Eddins, "Analysis/synthesis techniques for subband image coding", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol.38, No.8, pp.1446-1456, August 1990.
- [80] T.A. Standish, *Data Structures, Algorithms and Software Principles in C*, Addison Wesley, 1994.
- [81] D.B.H. Tay, "Design of filter banks using transformation of variables: new results", *IEEE Transactions on Signal Processing*, Vol.46, No.1, pp.203-209, January 1998.
- [82] B.H. Tay and N.G. Kingsbury, "Flexible design of multidimensional perfect reconstruction FIR 2-band filters using transformations of variables", *IEEE Transactions on Image Processing*, Vol.2, pp.466-480, October 1993.
- [83] B.H. Tay and N.G. Kingsbury, "Design of nonseparable 3-D filter banks/wavelet bases using transformations of variables", *IEE Proceeding: vision, image, and signal processing*, Vol.143, No.1, pp.51-61, February 1996,
- [84] T.D. Tran and T.Q. Nguyen, "On M -channel linear phase FIR filter banks and application in image compression", *IEEE Transactions on Signal Processing*, Vol.45, No.9, pp.2175-2187, September 1997.

- [85] M.K. Tsatsanis, G.B. Giannakis, "Multirate filter banks for code-division multiple access systems", *1995 IEEE ICASSP*, Vol.2, pp.1484-1487, 1995.
- [86] T.E. Tuncer and T.Q. Nguyen, "General analysis of two-band QMF banks", *IEEE Transactions on Signal Processing*, Vol.43, No.2, pp.544-548, February 1995.
- [87] P.P. Vaidyanathan, "Theory and design of M -channel maximally decimated quadrature mirror filters with arbitrary M , having perfect reconstruction property", *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol.ASSP-35, pp.476-492, April 1987.
- [88] P.P. Vaidyanathan, "Quadrature mirror filter banks, M -band extensions and perfect reconstruction techniques", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol.4, pp. 4-20, July 1987.
- [89] P.P. Vaidyanathan and P.-Q. Hoang, "Lattice structures for optimal design and robust implementation of two-channel perfect reconstruction QMF banks", *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol.36, pp.81-94, January 1988.
- [90] P.P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice-Hall, Englewood Cliffs, New Jersey, 1993.
- [91] R.N.J. Veldhuis, M. Breeuwer, and R.G. Van der Waal, "Subband coding of digital audio signals", *Philips Journal of Research*, Vol.44(2/3), pp. 329-343, 1989.
- [92] M. Vetterli, "Multidimensional subband coding: some theory and algorithms", *Signal Processing*, Vol.6, pp. 97-112, April 1984.
- [93] M. Vetterli, "A theory of multirate filter banks", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol.35, No.3, pp. 356-372, March 1987.
- [94] M. Vetterli and D.J. LeGall, "Perfect reconstruction FIR filter banks: some properties and factorizations", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol.37, No.7, pp. 1057-1071, July 1989.
- [95] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*, Prentice Hall PTR, Englewood Cliffs, New Jersey 07632, 1995.
- [96] S. Wada, "Design of nonuniform division multirate FIR filter banks", *IEEE Transactions on Circuits and Systems II*, Vol.42, No.2, pp.115-121, February 1995.

- [97] D. Wei, B.L. Evans, and A.C. Bovik, "Loss of perfect reconstruction in multidimensional filterbanks and wavelets designed via extended McClellan transformations", *IEEE Signal Processing Letters*, Vol.4, No.10, pp.295-297, October 1997.
- [98] J.W. Woods and S.D. O'Neil, "Subband coding of images", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol.ASSP-34, No.5, pp.1278-1288, October 1986.
- [99] J.W. Woods, *Subband Image Coding*, Norwell, MA: Kluwer, 1991.
- [100] G.W. Wornell, "Emerging applications of multirate signal processing and wavelets in digital communications", *Proceedings of IEEE*, Vol.84, pp.586-603, April 1996.
- [101] X.-G. Xia, "New precoding for intersymbol interference cancellation using nonmaximally decimated multirate filterbanks with ideal FIR equalizers", *IEEE Transactions on Signal Processing*, Vol.45, No.10, pp.2431-2441, October 1997.
- [102] H. Xu, W.-S. Lu, and A. Antoniou, "An efficient algorithm for the design of multi-channel cosine-modulated QMF banks", *1995 IEEE 38th Midwest Symposium on Circuits and Systems*, Vol.2, pp.1106-1109.
- [103] H. Xu, W.-S. Lu, and A. Antoniou, "Improved iterative methods for the design of quadrature mirror-image filter banks", *IEEE Transactions on Circuits and Systems II*, Vol. 43, pp.363-371, May 1996.
- [104] K. Zhou, J.C. Doyle and K. Glover, *Robust and Optimal Control*, Prentice-Hall, Upper Saddle River, New Jersey, 1996.

VITA

Jian Huang was born on November 22, 1967, in Shanghai, The People's Republic of China. After graduation from Shanghai Nan Yang Middle School in 1986, he studied in Shanghai University where he obtained a bachelor of science degree with honors in Electrical Engineering in 1990. Before long he became an electrical engineer and would have been spent the rest of his life in Shanghai, just like most ordinary people in that famous city. But because of his interest in scientific research, and great support from his family members and relatives across the Taiwan Strait, Jian later came to the United States and obtained a master of science degree in Electrical Engineering from Louisiana State University (LSU) in 1994. Afterwards, he was with LSU's Remote Sensing and Image Processing Laboratory for one more year. In Summer 1995, he made his decision to pursue a doctoral degree in the area of multirate digital signal processing. He worked as teaching assistant and research assistant during his graduate study from 1992 to 1998. He is presently a candidate for the degree of Doctor of Philosophy in Electrical Engineering at Louisiana State University.

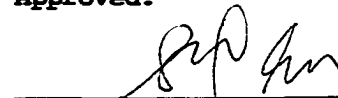
DOCTORAL EXAMINATION AND DISSERTATION REPORT

Candidate: Jian Huang

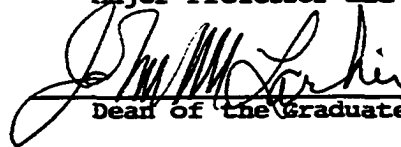
Major Field: Electrical Engineering

Title of Dissertation: A System Approach to the Design of Multirate
Filter Banks

Approved:

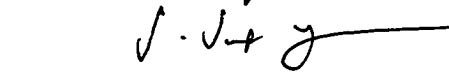
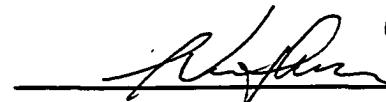
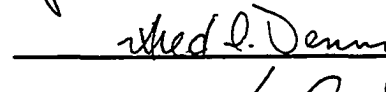
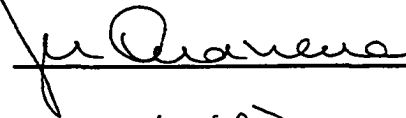


Major Professor and Chairman



Dean of the Graduate School

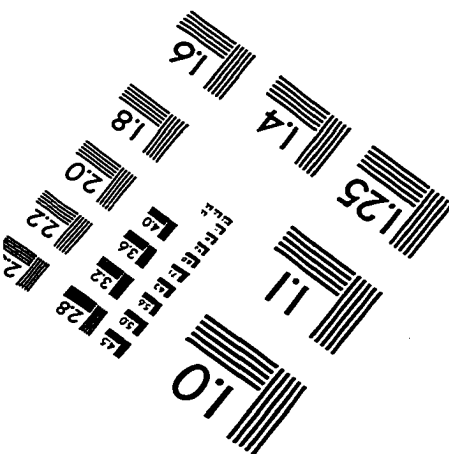
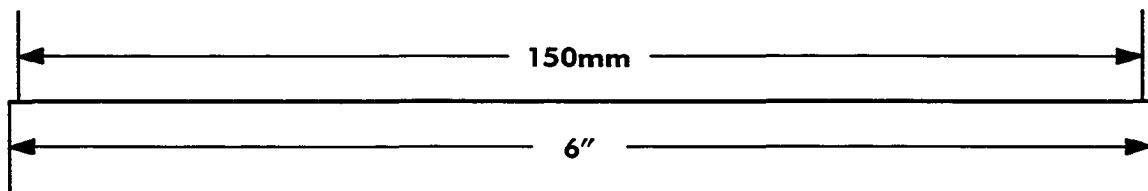
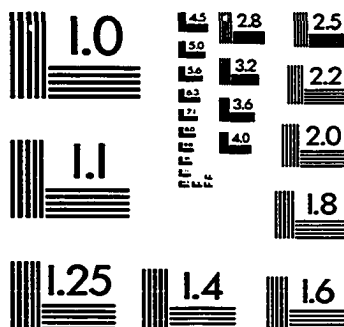
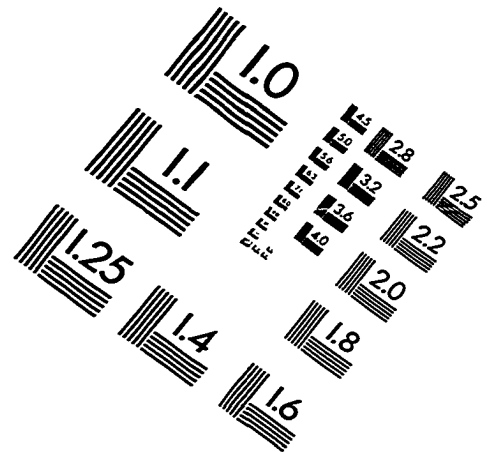
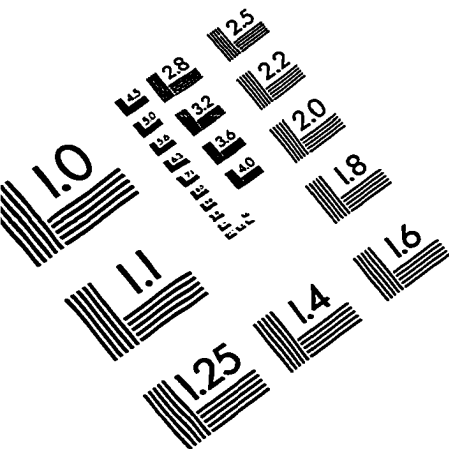
EXAMINING COMMITTEE:



Date of Examination:

June 11, 1998

IMAGE EVALUATION TEST TARGET (QA-3)



APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved

